

181 CAGATCTGGATGACCCAGAGGACATTGGAATTCACAAAGACCTCGTGGAAAGTG 240
241 CTGCTGGATGCACTTTGGCAACCACTCTACCTCTGTGAAAGTGACTATTTTGGCTCGAG 300
241 CTGCTGGATGCACTTTGGCAACCACTCTACCTGTGAAAGTGACTATTTTGGCTCGAG 300
301 TTTCTGATCACAAAAGATCAGGTGTGGCTGATCTCTTAAACCAATGTGAAACAG 360
301 TTTCTGATCACAAAAGATCAGGTGTGGCTGATCTCTTAAACCAATGTGAAACAG 360
361 ATTAGAAGCCAAAGACGTTGTGTAAATTGTGTAAATTCTTTCCGCTGACAC 420
361 ATTAGAAGCCAAAGACGTTGTGTAAATTGTGTAAATTCTTTCCGCTGACAC 420
421 AACACATCTCCAGAAAGAACTCACAAAGTACCTGTGGCGCTGAGGTGAAGCAGACTTG 480
421 AACACATCTCCAGAAAGAACTCACAAAGTACCTGTGGCGCTGAGGTGAAGCAGACTTG 480
481 GCTCAAGGAGGTGACGTGTAAATGACACAGCGACGCTCTTGATTTCAACATTTGTG 540
481 GCTCAAGGAGGTGACGTGTAAATGACACAGCGACGCTCTTGATTTCAACATTTGTG 540
541 CAATCTGAGATTGGGATTTTGAATGAAAGCTTGGACAGAGACCTTACCAAAAAATAA 600
541 CAATCTGAGATTGGGATTTTGAATGAAAGCTTGGACAGAGACCTTACCAAAAAATAA 600
601 TACATCTCTCAGACAGACCACTAGAGACCAAAATGTGTGAATTTTCAACATAACCAT 660
601 TACATCTCTCAGACAGACCACTAGAGACCAAAATGTGTGAATTTTCAACATAACCAT 660
661 GGAACAAACACAGACAAATCAGATTTCCAGCTCTAGAGATTGCCGTGGCTAGAGATG 720
661 GGAACAAACACAGACAAATCAGATTTCCAGCTCTAGAGATTGCCGTGGCTAGAGATG 720
721 TATGAATCCGGTTGACCCGCGCAAGAGACAGGAAAGGACAGAAATCAATCTGCGGTT 780
721 TATGAATCCGGTTGACCCGCGCAAGAGACAGGAAAGGACAGAAATCAATCTGCGGTT 780
781 GCCAACAACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCCCTTCAACTGGGCG 840
781 GCCAACAACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCCCTTCAACTGGGCG 840
841 AAGGTGCGAAGGTGAGTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGGCCAGATGCC 900
841 AAGGTGCGAAGGTGAGTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGGCCAGATGCC 900
901 AATAGTGCCTACAGGATACCTTGAATTCCTGATGCGCAGTCGGGATTTCTGCAAGTCC 960
901 AATAGTGCCTACAGGATACCTTGAATTCCTGATGCGCAGTCGGGATTTCTGCAAGTCC 960
961 TTTCTGAAAATCTGTGTGAACATCAATGCCCTTCTTGAAGTCTTTTGAAGGCCAAACCA 1020
961 TTTCTGAAAATCTGTGTGAACATCAATGCCCTTCTTGAAGTCTTTTGAAGGCCAAACCA 1020
1021 AAGCCCAAGCCCGTCTCTTTAAGCCGCGGATCATCTTCCGTTCACTGATGCGACTCAG 1080
1021 AAGCCCAAGCCCGTCTCTTTAAGCCGCGGATCATCTTCCGTTCACTGATGCGACTCAG 1080
1081 AAGCAGGTTCTGACATGTATTAAGAGAGGACATTAAGAGGTGCACTTTGAAGGAAG 1140
1081 AAGCAGGTTCTGACATGTATTAAGAGAGGACATTAAGAGGTGCACTTTGAAGGAAG 1140
1141 CACAGAAATTCATCTATCCGAGACCTTGTCTTCAAGCTTACAAATCGAATTTGGAA 1200
1141 CACAGAAATTCATCTATCCGAGACCTTGTCTTCAAGCTTACAAATCGAATTTGGAA 1200
1201 GTGCTGAGACAGTCTCAGCAGAGACACAGCTTACATTTGAGAGAGGTGCCGATCTCA 1260
1201 GTGCTGAGACAGTCTCAGCAGAGACACAGCTTACATTTGAGAGAGGTGCCGATCTCA 1260
1261 GGGGAGCAGAGCTGCGGCGAGAGAAAGAACGAAAGTTTCGCGCGGAGACCGGAGTCTG 1320
1261 GGGGAGCAGAGCTGCGGCGAGAGAAAGAACGAAAGTTTCGCGCGGAGACCGGAGTCTG 1320

1261 GGGGAGCAGAGCTGCGGCGAGAGAAAGAACGAAAGTTTCGCGCGGAGACCGGAGTCTG 1320
1321 CACCCGAGCCCTGCGCCGAGAGAAAGCCCGCGGTTAAACAGCAGCGAGCGAGCCGCG 1380
1321 CACCCGAGCCCTGCGCCGAGAGAAAGCCCGCGGTTAAACAGCAGCGAGCGAGCCGCG 1380
1381 TCGGCGCCACAGAGAGAGAGAGAGGTCTGTTAAGATAGAGACCAGCAGATTAACCT 1440
1381 TCGGCGCCACAGAGAGAGAGAGAGGTCTGTTAAGATAGAGACCAGCAGATTAACCT 1440
1441 CAGCCCCGAGCCAGACACAGAGCTCTCTGACTGAGCTCTCACTTTCCGAGCTGTCT 1500
1441 CAGCCCCGAGCCAGACACAGAGCTCTCTGACTGAGCTCTCACTTTCCGAGCTGTCT 1500
1501 GTGAAGTGGAGGGGAGGTGGCCCTGACAGGACCTTGTCTCCCAACCTGAGCCCG 1560
1501 GTGAAGTGGAGGGGAGGTGGCCCTGACAGGACCTTGTCTCCCAACCTGAGCCCG 1560
1561 GACACCAAGCAGGCTCTCTCTGATCAGCCGCTGTAATGACAGGCTGCCCCGG 1620
1561 GACACCAAGCAGGCTCTCTCTGATCAGCCGCTGTAATGACAGGCTGCCCCGG 1620
1621 ACAGACATGAGATGAGGGCCGAGAGAGATTCCTCACTGATTAAGCTTACTATA 1680
1621 ACAGACATGAGATGAGGGCCGAGAGAGATTCCTCACTGATTAAGCTTACTATA 1680
1681 GCTAAGAGAGTCTACCAACGAGGAAATATCTGAAGATCTGAAAGTTTCACTTCTG 1740
1681 GCTAAGAGAGTCTACCAACGAGGAAATATCTGAAGATCTGAAAGTTTCACTTCTG 1740
1741 TGGTTTCAGAGCAGTGAAGCAAGAGAGCGCATGCCGGAAGCATGAAAAGTCTCAT 1800
1741 TGGTTTCAGAGCAGTGAAGCAAGAGAGCGCATGCCGGAAGCATGAAAAGTCTCAT 1800
1801 TTCCCGAATTTTGAACCTTTTGAACAAATTTTCAATATTTTCTCAAGAAATTAAGCA 1860
1801 TTCCCGAATTTTGAACCTTTTGAACAAATTTTCAATATTTTCTCAAGAAATTAAGCA 1860
1861 GCACTTGCCCTGTGGAGAGCGCGCTCAATGCGCCAAATCAGAGATTACAAAGAAATCGGC 1920
1861 GCACTTGCCCTGTGGAGAGCGCGCTCAATGCGCCAAATCAGAGATTACAAAGAAATCGGC 1920
1921 GATGTCATGCTGAAGAACATTCAGGGCATGAAGCACTGCGGCTCACCTGTGAAGCAC 1980
1921 GATGTCATGCTGAAGAACATTCAGGGCATGAAGCACTGCGGCTCACCTGTGAAGCAC 1980
1981 AGCGAGGCTTGGAGGCCCTGAGAAATGAAGTCAAGACTCCCGGCGCTGAGAACTTC 2040
1981 AGCGAGGCTTGGAGGCCCTGAGAAATGAAGTCAAGACTCCCGGCGCTGAGAACTTC 2040
2041 TGCAGAGCTTTGAGCTCAGAAAGGTGTGTTACTACCGCTCAACACTTCTCTGCGG 2100
2041 TGCAGAGCTTTGAGCTCAGAAAGGTGTGTTACTACCGCTCAACACTTCTCTGCGG 2100
2101 CCACTGACCGGCTGATCACTACAGAGGTCTTGAAGCGGCTGTGCAACACCAACCG 2160
2101 CCACTGACCGGCTGATCACTACAGAGGTCTTGAAGCGGCTGTGCAACACCAACCG 2160
2161 CCGAGCCACGCGCACTTTAGAGGACTGCGCAGACCGCTTTTGGCAGAGATCAAGAGATGG 2220
2161 CCGAGCCACGCGCACTTTAGAGGACTGCGCAGACCGCTTTTGGCAGAGATCAAGAGATGG 2220
2221 GCACAGCTCCACGAGATGATCAAGATGAGAAATTTCCAGAGCTGCAAGACTCAAG 2280
2221 GCACAGCTCCACGAGATGATCAAGATGAGAAATTTCCAGAGCTGCAAGACTCAAG 2280
2281 AAAAATTTGATTTGCAATCTTGTGTTCCGAGAGGAGATTCATCGTCTGAGC 2340
2281 AAAAATTTGATTTGCAATCTTGTGTTCCGAGAGGAGATTCATCGTCTGAGC 2340
2341 AGCTTCAGCAAGCTTCCGAGAGGAGGCTCCAGAGCGCATGTTCTTCTGTTCAAGAC 2400
2341 AGCTTCAGCAAGCTTCCGAGAGGAGGCTCCAGAGCGCATGTTCTTCTGTTCAAGAC 2400
2441 AGCTTCAGCAAGCTTCCGAGAGGAGGCTCCAGAGCGCATGTTCTTCTGTTCAAGAC 2400
2441 AGCTTCAGCAAGCTTCCGAGAGGAGGCTCCAGAGCGCATGTTCTTCTGTTCAAGAC 2400

|||||
1 CCGCGAGCGCGCGCTGTGGAGATTTCTTAAGCGCTTTCATGAGGAGAAATA 60
QY 61 GAGCGAAGGCCGACCCCGAGATCAAGATCGGGGGCCCCGGAAAAATTCGGGGATCACTAC 120
Db 61 GAGCGAAGGCCGACCCCGAGATCAAGATCGGGGGCCCCGGAAAAATTCGGGGATCACTAC 120
QY 121 TTGGAACTGTGACAGAAAGCCGCCCAACACTTCAGAAAACTCGTGTCAATCAAAATC 180
Db 121 TTGGAACTGTGACAGAAAGCCGCCCAACACTTCAGAAAACTCGTGTCAATCAAAATC 180
QY 181 CAGATCTGTGATGACCCGAGAGGCAATTTGAAGTTTCAAAAGACTCTCTGGAAAGTG 240
Db 181 CAGATCTGTGATGACCCGAGAGGCAATTTGAAGTTTCAAAAGACTCTCTGGAAAGTG 240
QY 241 CTGCTGGAGTGTGCAACCACTCAACCTGTGGAAAGTGACATTTTGGCCCTCGAG 300
Db 241 CTGCTGGAGTGTGCAACCACTCAACCTGTGGAAAGTGACATTTTGGCCCTCGAG 300
QY 301 TTTCTGATCACAAAAAGATCAAGGTGTGGCTGATCTCTTAAACCAATTTGAACAG 360
Db 301 TTTCTGATCACAAAAAGATCAAGGTGTGGCTGATCTCTTAAACCAATTTGAACAG 360
QY 361 ATTGAAGGCCAAAGCAGTTGTTTAAAGTTGTGTGAATTTCTTTCGCTGACAC 420
Db 361 ATTGAAGGCCAAAGCAGTTGTTTAAAGTTGTGTGAATTTCTTTCGCTGACAC 420
QY 421 AACAACTCCAGAAAGATCAAGATCAAGTACCTGTGCGCTGAGGTGAAGCAGACTTG 480
Db 421 AACAACTCCAGAAAGATCAAGATCAAGTACCTGTGCGCTGAGGTGAAGCAGACTTG 480
QY 481 GCTCAAGGAGGTGTGACGTGTAATGACACGAGCGCAGCTCTCTTGAATTCACATTTG 540
Db 481 GCTCAAGGAGGTGTGACGTGTAATGACACGAGCGCAGCTCTCTTGAATTCACATTTG 540
QY 541 CAATCTGAGATTTGGGATTTTGTATGAAAGCTTTGACAGAGAGCATTTAGCAAAAATTA 600
Db 541 CAATCTGAGATTTGGGATTTTGTATGAAAGCTTTGACAGAGAGCATTTAGCAAAAATTA 600
QY 601 TACATCTCAGAGAAAGCAGCATGAGAGCAAAATGTGGAAATTTACATTAACCACTT 660
Db 601 TACATCTCAGAGAAAGCAGCATGAGAGCAAAATGTGGAAATTTACATTAACCACTT 660
QY 661 GGAACAACACAGAGAAATCAGATTTCAAGCTCTTGAAGTTGCCGCTGAGAGATG 720
Db 661 GGAACAACACAGAGAAATCAGATTTCAAGCTCTTGAAGTTGCCGCTGAGAGATG 720
QY 721 TATGAATCCGTTGACCCGCGCAAGAGCAGGGAAGGCAAGAAATCAATCTGCCGTT 780
Db 721 TATGAATCCGTTGACCCGCGCAAGAGCAGGGAAGGCAAGAAATCAATCTGCCGTT 780
QY 781 GCCAAACGAGGAATCTAAGTTTCAAGGTTTCAATGAATCAATGCTTCAACTGCGGC 840
Db 781 GCCAAACGAGGAATCTAAGTTTCAAGGTTTCAATGAATCAATGCTTCAACTGCGGC 840
QY 841 AAGGTGCGAAGCTGAGCTTCAAGAGAAAGCTTTCTCATCAAGCTCCGCGCAGATGCG 900
Db 841 AAGGTGCGAAGCTGAGCTTCAAGAGAAAGCTTTCTCATCAAGCTCCGCGCAGATGCG 900
QY 901 AATATGCGTACAGAGATCTTTGGAATTCGTATGCGCAGTGGGATTTCTGCAAGTCC 960
Db 901 AATATGCGTACAGAGATCTTTGGAATTCGTATGCGCAGTGGGATTTCTGCAAGTCC 960
QY 961 TTTCTGAAAATCTGTGTGAACATCAATGCTTCTTGAATTTTGAAGGCCCAACCA 1020
Db 961 TTTCTGAAAATCTGTGTGAACATCAATGCTTCTTGAATTTTGAAGGCCCAACCA 1020
QY 1021 AAGCCCAAGCCGCTCTCTTAAAGCCGAGGATCAATTCGATTCAGTGTGCACTCAG 1080
Db 1021 AAGCCCAAGCCGCTCTCTTAAAGCCGAGGATCAATTCGATTCAGTGTGCACTCAG 1080
QY 1081 AAGCAGTTCTGACATATGTTAAAGAGAGCATTAAGAGGTGACATTTGAAGGAAG 1140
|||||

Db 1081 AAGCAGTTCTGACATATGTTAAAGAGAGCATTAAGAGGTGACATTTGAAGGAAG 1140
QY 1141 CACAGCAAGATTCATTTATTCAGAGAGCTTGTCTTCAAGCTTACAGAACTGAATTCGGA 1200
Db 1141 CACAGCAAGATTCATTTATTCAGAGAGCTTGTCTTCAAGCTTACAGAACTGAATTCGGA 1200
QY 1201 GTGCTGAGCAGTCTCAGCAGAGCAGAGCTTCAATTTGAGAAAGGTGCGGAATCTCCA 1260
Db 1201 GTGCTGAGCAGTCTCAGCAGAGCAGAGCTTCAATTTGAGAAAGGTGCGGAATCTCCA 1260
QY 1261 GGGGGCCAGAGTGTGCGCGCAGAGAAAGAACCGAAGTTTCCGCGGGGAGGCGGGGTCG 1320
Db 1261 GGGGGCCAGAGTGTGCGCGCAGAGAAAGAACCGAAGTTTCCGCGGGGAGGCGGGGTCG 1320
QY 1321 CACCGGAGCCCTGCGCGCAGAGAGAAAGCCCGCGGGTTAAAGAGAGGCGAGCGAGCC 1380
Db 1321 CACCGGAGCCCTGCGCGCAGAGAGAAAGCCCGCGGGTTAAAGAGAGGCGAGCGAGCC 1380
QY 1381 TCGGCGCCCAAGAGAGAAAGAGAGAGGTCTGTTAAGATTAAGACCAGCAGATTAACCT 1440
Db 1381 TCGGCGCCCAAGAGAGAAAGAGAGAGGTCTGTTAAGATTAAGACCAGCAGATTAACCT 1440
QY 1441 CAGCCCCCGCAGCAGAGCAGAGCTCCCTGACTGCACTCTTCACTTCCGAGGCTGCT 1500
Db 1441 CAGCCCCCGCAGCAGAGCAGAGCTCCCTGACTGCACTCTTCCGAGGCTGCT 1500
QY 1501 GTGAAGTGGAGGGGAGAGTGGCCCTGCAAGCTGACCTTGTCTCCCAACTGAGCCCC 1560
Db 1501 GTGAAGTGGAGGGGAGAGTGGCCCTGCAAGCTGACCTTGTCTCCCAACTGAGCCCC 1560
QY 1561 GACACCAAGCAGGCTCTCTCTTATCAAGCCGCTGTGATGACAGAGCTGCCCCGG 1620
Db 1561 GACACCAAGCAGGCTCTCTCTTATCAAGCCGCTGTGATGACAGAGCTGCCCCGG 1620
QY 1621 ACGAGCATGAGAGAGAGAGGCGCGAGAGAGATTTCCCACTGAATTAAGCGATTCATA 1680
Db 1621 ACGAGCATGAGAGAGAGAGGCGCGAGAGAGATTTCCCACTGAATTAAGCGATTCATA 1680
QY 1681 GCTAAGAAAGTGTCTACCAACGAGCGAACTATCTGAAGATCTCGAAGTTTCACTTGC 1740
Db 1681 GCTAAGAAAGTGTCTACCAACGAGCGAACTATCTGAAGATCTCGAAGTTTCACTTGC 1740
QY 1741 TGGTTTCAAGAGCAGTGAAGCAAGAGAGCAGCATGCGGAAAGCACTGAAAAGTCTCAT 1800
Db 1741 TGGTTTCAAGAGCAGTGAAGCAAGAGAGCAGCATGCGGAAAGCACTGAAAAGTCTCAT 1800
QY 1801 TTCCGGAATTTGAACCTTTGACAAATTTCTAATAATTTTCTCAAGAAATTAAGCAA 1860
Db 1801 TTCCGGAATTTGAACCTTTGACAAATTTCTAATAATTTTCTCAAGAAATTAAGCAA 1860
QY 1861 CGACTTGCCTGTGGGAAGGCGCTCAATGCGCAATTCAGAGATTAACCAAGAAATCGGC 1920
Db 1861 CGACTTGCCTGTGGGAAGGCGCTCAATGCGCAATTCAGAGATTAACCAAGAAATCGGC 1920
QY 1921 GATGTATGCTGAAGAAATTCAGAGGCAATGAAAGCACTGCGCGCTCACTGTGGAAGCAC 1980
Db 1921 GATGTATGCTGAAGAAATTCAGAGGCAATGAAAGCACTGCGCGCTCACTGTGGAAGCAC 1980
QY 1981 AGCGAGGCTTGAAGAGCCCTGGAAGATGAATCAAGAGCTCCCGCGGCTGGAAGAACTTC 2040
Db 1981 AGCGAGGCTTGAAGAGCCCTGGAAGATGAATCAAGAGCTCCCGCGGCTGGAAGAACTTC 2040
QY 2041 TGCAGAGACTTGAAGCTCAGAAAGGTGTGTTACTTACCGCTCAACACTTCTCTCGGCG 2100
Db 2041 TGCAGAGACTTGAAGCTCAGAAAGGTGTGTTACTTACCGCTCAACACTTCTCTCGGCG 2100
QY 2101 CCACTGACCGGCTCATGCACTAAGAGAGGCTCTGAGAGCGCTGTGCAACACACCGG 2160
Db 2101 CCACTGACCGGCTCATGCACTAAGAGAGGCTCTGAGAGCGCTGTGCAACACACCGG 2160
QY 2161 CCGAGCCACGCGCACTTGAAGGAGCTGCGAGCGCTTTTGGCAGAGATCAAGAGATGAGT 2220
Db 2161 CCGAGCCACGCGCACTTGAAGGAGCTGCGAGCGCTTTTGGCAGAGATCAAGAGATGAGT 2220

QY 2221 GCACAGCTCCAGGGTACGATGATCAAGTGGAGAAATTTCCAGAAAGCTGCAAGAACTCAAG 2280
 DB 2221 GCACAGCTCCAGGGTACGATGATCAAGTGGAGAAATTTCCAGAAAGCTGCAAGAACTCAAG 2280
 QY 2281 AAGATTTGATTTGGATTTGACAAATCTTGTGGTCCGGGAAAGGAGTTCACTCCCTGAGG 2340
 DB 2281 AAGATTTGATTTGGATTTGACAAATCTTGTGGTCCGGGAAAGGAGTTCACTCCCTGAGG 2340
 QY 2341 AGCTCAGCAAGCTCTCGGGGAAAGGGGCTCAGAGAGCATGTTCTTCTGTTCAAGAC 2400
 DB 2341 AGCTCAGCAAGCTCTCGGGGAAAGGGGCTCAGAGAGCATGTTCTTCTGTTCAAGAC 2400
 QY 2401 GTCTCTGATTAACAGAGCGGGGGCTGAGAGGCTCCAAATAGTTTAAAGTCCAGGGGAG 2460
 DB 2401 GTCTCTGATTAACAGAGCGGGGGCTGAGAGGCTCCAAATAGTTTAAAGTCCAGGGGAG 2460
 QY 2461 CTCCCGCTCTATGCGATGACGATTTGAGAGAGCGAAGACGAGTGGGGGGTCCCACTGC 2520
 DB 2461 CTCCCGCTCTATGCGATGACGATTTGAGAGAGCGAAGACGAGTGGGGGGTCCCACTGC 2520
 QY 2521 CTGACCTTCGGGGGCGAGCGGCGATCCATCTGTGGCGGCACTTCTGGTCCGAGATG 2580
 DB 2521 CTGACCTTCGGGGGCGAGCGGCGATCCATCTGTGGCGGCACTTCTGGTCCGAGATG 2580
 QY 2581 GAGAAAGTGGTTGAGAGACATCCAGATGGCCATTTGACCTGGCGGAGAGAGACAGACCCC 2640
 DB 2581 GAGAAAGTGGTTGAGAGACATCCAGATGGCCATTTGACCTGGCGGAGAGAGACAGACCCC 2640
 QY 2641 GCCCCTGAGTCTCTGCGCAGACAGCCCCCTGACAAAGTCCCTGATGAGGACACCGG 2700
 DB 2641 GCCCCTGAGTCTCTGCGCAGACAGCCCCCTGACAAAGTCCCTGATGAGGACACCGG 2700
 QY 2701 GCTGACCAAGAGTCAAGAGATGACTGAGCGCTTCGCGCATCTGCTGAGAGCGCCAGGCC 2760
 DB 2701 GCTGACCAAGAGTCAAGAGATGACTGAGCGCTTCGCGCATCTGCTGAGAGCGCCAGGCC 2760
 QY 2761 CGGACCGGGGGGCAACATGATGTCAGTGTGGGCAACGGAACACGAGCGTCCATG 2820
 DB 2761 CGGACCGGGGGGCAACATGATGTCAGTGTGGGCAACGGAACACGAGCGTCCATG 2820
 QY 2821 GTGAGATTGACGATGCGACGTGAGAGATCACTGTCTGAGAACTGCTGAGGAAATTCAAA 2880
 DB 2821 GTGAGATTGACGATGCGACGTGAGAGATCACTGTCTGAGAACTGCTGAGGAAATTCAAA 2880
 QY 2881 AACAGCAAGGCTGCGAGAACTGTGGGTGTTCACAACTTCTGCTGTTCTTAC 2940
 DB 2881 AACAGCAAGGCTGCGAGAACTGTGGGTGTTCACAACTTCTGCTGTTCTTAC 2940
 QY 2941 AATTCACACAGGACATATCCCTTTCAGAGCTGCTGCTGCGGTAATGCTCAC 3000
 DB 2941 AATTCACACAGGACATATCCCTTTCAGAGCTGCTGCTGCGGTAATGCTCAC 3000
 QY 3001 ATCCCTCTGAGTCCGAGAACTCAAGAAAGTCAAGTGTTCAGAGTCACTCAAGTCC 3060
 DB 3001 ATCCCTCTGAGTCCGAGAACTCAAGAAAGTCAAGTGTTCAGAGTCACTCAAGTCC 3060
 QY 3061 CACGTCTACTACTTCAAGGCGGAAAGCGAGTACAGTTGAAAGGTGATGAAAGTATC 3120
 DB 3061 CACGTCTACTACTTCAAGGCGGAAAGCGAGTACAGTTGAAAGGTGATGAAAGTATC 3120
 QY 3121 CGCAGTCCACAGGCTCTGCTGCGACCCCAAGTGTGAGCCACAAAGAGTCTCTGAG 3180
 DB 3121 CGCAGTCCACAGGCTCTGCTGCGACCCCAAGTGTGAGCCACAAAGAGTCTCTGAG 3180
 QY 3181 TATTGATGCGCGGACACCTGTTTCCGAGTGGCTTCTTCTGAGAGAGCTTCTCTT 3240
 DB 3181 TATTGATGCGCGGACACCTGTTTCCGAGTGGCTTCTTCTGAGAGAGCTTCTCTT 3240
 QY 3241 CTCTGTATTAAAGAGCTGGTAAATTAACCTGTCTGAAATCAAAAATCATGGCTT 3300
 DB 3241 CTCTGTATTAAAGAGCTGGTAAATTAACCTGTCTGAAATCAAAAATCATGGCTT 3300

QY 3301 CCAGAGCTCTCTGTCTCCACAGCGGTTTTTAAACCCGAGCTCTCAGCGTTTGA 3360
 DB 3301 CCAGAGCTCTCTGTCTCCACAGCGGTTTTTAAACCCGAGCTCTCAGCGTTTGA 3360
 QY 3361 TGAACAGCGCTCCACCTTCAGTCTGCGATCGCTGGGGGCGGTGTTCTTAAGTATG 3420
 DB 3361 TGAACAGCGCTCCACCTTCAGTCTGCGATCGCTGGGGGCGGTGTTCTTAAGTATG 3420
 QY 3421 CCAGTATTAAACATTGTCAAT 3442
 DB 3421 CCAGTATTAAACATTGTCAAT 3442
 RESULT 3
 CS130723
 LOCUS CS130723 5028 bp DNA linear PAT 02-AUG-2005
 DEFINITION Sequence 9 from Patent WO2005064009.
 ACCESSION CS130723
 VERSION CS130723.1 GI:71792793
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE 1
 AUTHORS Oernboft, T. F.
 TITLES Classification of cancer
 JOURNAL Patent: WO 2005064009-A 9 14-JUL-2005;
 Aros Applied Biotechnology APS (DK)
 FEATURES
 source 1..5028
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="FERM, RhoGEP (ARHGEF) and plectstrin domain pro"
 ORIGIN
 Query Match 99.9%; Score 3438.8; DB 2; Length 5028;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3440; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CGCCGACGCGCGCGGCTGTGAGATATTTCTAAGCCGTTTCAATGAGAGAAAT 60
 DB 289 CGCCGACGCGCGCGGCTGTGAGATATTTCTAAGCCGTTTCAATGAGAGAAAT 348
 QY 61 GAGCAGAGGCGGACCCGAGGATCAAGCTGGGGGCCCGGAAATTCGGGGATCACTACC 120
 DB 349 GAGCAGAGGCGGACCCGAGGATCAAGCTGGGGGCCCGGAAATTCGGGGATCACTACC 408
 QY 121 TTGGAACGTGACAGAAAGCGGCCCCCAACCTTCAGAGAAATCTGTGTCATCAAAATC 180
 DB 409 TTGGAACGTGACAGAAAGCGGCCCCCAACCTTCAGAGAAATCTGTGTCATCAAAATC 468
 QY 181 CAGATGCTGATGACACCCAGAGAGGATTTGAAGTTCCACAAAGAGCTCTGGGAAGGTG 240
 DB 469 CAGATGCTGATGACACCCAGAGAGGATTTGAAGTTCCACAAAGAGCTCTGGGAAGGTG 528
 QY 241 CTGCTGATGACAGTTTGAACCACTCAACTGTGGAAGGTGATATTTTGGCTCGAG 300
 DB 529 CTGCTGATGACAGTTTGAACCACTCAACTGTGGAAGGTGATATTTTGGCTCGAG 588
 QY 301 TTTCTGATCACAAAAGATCAAGGTGAGTGTGATCTCTTAAACCTATGGAACAG 360
 DB 589 TTTCTGATCACAAAAGATCAAGGTGAGTGTGATCTCTTAAACCTATGGAACAG 648
 QY 361 ATTAGAAGCCAAAGACCTGTTGTTAAGTTTGAGTAAATTTCTTCCGCTGACAC 420
 DB 649 ATTAGAAGCCAAAGACCTGTTGTTAAGTTTGAGTAAATTTCTTCCGCTGACAC 708
 QY 421 ACACAACTCCAAAGAACTCAACAAAGTACTGTTGCGCTGCAAGTGAAGACGACTTG 480
 DB 709 ACACAACTCCAAAGAACTCAACAAAGTACTGTTGCGCTGCAAGTGAAGACGACTTG 768

481 GCTCAAGGAGGTTGACGTATATGACACAGGAGCTCTCTTGATTTACACATTTGG 540
482 GCTCAAGGAGGTTGACGTATATGACACAGGAGCTCTCTTGATTTACACATTTGG 828
541 CAATCGAGATTGGGATTTTGAATGAAGCTTGAACAGAGCACTTACCAAAATATA 600
829 CAATCGAGATTGGGATTTTGAATGAAGCTTGAACAGAGCACTTACCAAAATATA 888
601 TACATCTCTCAGACAGACCACTAGAGGACCAAAATGTGGAATTTACATACCACTT 660
889 TACATACCTCAGACAGACCACTAGAGGACCAAAATGTGGAATTTACATACCACTT 948
661 GGAACAAACACAGAGCAATCAATTTCCAGCTCTTGAAGATTGCCGCGGCTAGAGATG 720
949 GGAACAAACACAGAGCAATCAATTTCCAGCTCTTGAAGATTGCCGCGGCTAGAGATG 1008
721 TATGGAATCCGGTTGACCCCGGCAAGGACAGGGAAGGACGAAATCAATCTGGCCGT 780
1009 TATGGAATCCGGTTGACCCCGGCAAGGACAGGGAAGGACGAAATCAATCTGGCCGT 1068
781 GCCAACACGGGAATTTCTAGTTTTCAGGGTTTCACTAATCAATGCTTCAACTGGGCC 840
1069 GCCAACACGGGAATTTCTAGTTTTCAGGGTTTCACTAATCAATGCTTCAACTGGGCC 1128
841 AAGGTCCGGAAGCTGAGCTTCAAGAGGAAGCGCTTCTCATCAAGCTCCGGCCAGATGCG 900
1129 AAGGTCCGGAAGCTGAGCTTCAAGAGGAAGCGCTTCTCATCAAGCTCCGGCCAGATGCG 1188
901 AATAGTCCGTACAGAGATCTTGAATTTCTGATGAGCAGATCGGATTTCTGCAAGTCC 960
1189 AATAGTCCGTACAGAGATCTTGAATTTCTGATGAGCAGATCGGATTTCTGCAAGTCC 1248
961 TTCTGGAATATCTGTGTTGAATCATCTGCTTCTTGAATTTTGAAGGCCCAACCA 1020
1249 TTCTGGAATATCTGTGTTGAATCATCTGCTTCTTGAATTTTGAAGGCCCAACCA 1308
1021 AAGCCCAAGCCGCTCTTTAGCCGGGGATCATCTTCCGTTCAAGTGGTGGATCTGAG 1080
1309 AAGCCCAAGCCGCTCTTTAGCCGGGGATCATCTTCCGTTCAAGTGGTGGATCTGAG 1368
1081 AAGCAGATTCTGACTATGTAAAGAGAGACATTAAGAGGTGACATTTGAAGGAAG 1140
1369 AAGCAGATTCTGACTATGTAAAGAGAGACATTAAGAGGTGACATTTGAAGGAAG 1428
1141 CACAGGAATATCTATCTATCCGAGACCTTGTCTTCAACCTTACAGAACTGAATTCGGA 1200
1429 CACAGGAATATCTATCTATCCGAGACCTTGTCTTCAACCTTACAGAACTGAATTCGGA 1488
1201 GTGCTGAGACAGTCTCAGAGAGACACAGCCTTACATTTGAGAGAGTCCGAAATCTCA 1260
1489 GTGCTGAGACAGTCTCAGAGAGACACAGCCTTACATTTGAGAGAGTCCGAAATCTCA 1548
1261 GGGGGGACAGAGCTGCGGCGAGAGAAAGAACCGAAGTTTCCGCGGGAGCCCGGGGTG 1320
1549 GGGGGGACAGAGCTGCGGCGAGAGAAAGAACCGAAGTTTCCGCGGGAGCCCGGGGTG 1608
1321 CACCCGAGGCCCTGCGCGGAGAGAAAGCCCGCGGGTAAACAAGAGCCGGAACCGAGCCG 1380
1609 CACCCGAGGCCCTGCGCGGAGAGAAAGCCCGCGGGTAAACAAGAGCCGGAACCGAGCCG 1668
1381 TCGGCGCCCAAGAGAGAGAGAGAGTCTGTTAAGATAGAACCCAGAGAGTAAACT 1440
1669 TCGGCGCCCAAGAGAGAGAGAGAGTCTGTTAAGATAGAACCCAGAGAGTAAACT 1728
1441 CAGCCCTCGACCAAGACAGAGCTCTGACTGAGAGTCTTCACTTTCCAGACTGTCT 1500
1729 CAGCCCTCGACCAAGACAGAGCTCTGACTGAGAGTCTTCACTTTCCAGACTGTCT 1788
1501 GTGAATCTCGAGAGAGAGAGTGGCCCTTGTCAACGTACCTTGTCTCCCAACTGAGCCCC 1560
1789 GTGAATCTCGAGAGAGAGAGTGGCCCTTGTCAACGTACCTTGTCTCCCAACTGAGCCCC 1848

1561 GACACCAAGAGGCTCTCCCTTGATCAGCCCGCTGTGAATGACACAGGCTGCCCCGG 1620
1849 GACACCAAGAGGCTCTCCCTTGATCAGCCCGCTGTGAATGACACAGGCTGCCCCGG 1908
1621 ACGAGCATGAGATGAGGCGCGAGAGAGATTTCCCACTGATTAAGCTTCTATTA 1680
1909 ACGAGCATGAGATGAGGCGCGAGAGAGATTTCCCACTGATTAAGCTTCTATTA 1968
1681 GCTAAGAGGTCTACACAGGAGGACATATCTGAAGGATCTCGAAGTTATCACTTCG 1740
1969 GCTAAGAGGTCTACACAGGAGGACATATCTGAAGGATCTCGAAGTTATCACTTCG 2028
1741 TGGTTTCAGAGCACAGTAGAGCAAGAGAGCCATGCGGAGAGCACTGAAAAGTCTCAT 1800
2029 TGGTTTCAGAGCACAGTAGAGCAAGAGAGCCATGCGGAGAGCACTGAAAAGTCTCAT 2088
1801 TTCCGGAATTTTGAACCTTTGACAAATTTTCACTAATTTTCTCAGAGAAATGAGCA 1860
2089 TTCCGGAATTTTGAACCTTTGACAAATTTTCACTAATTTTCTCAGAGAAATGAGCA 2148
1861 GCACTTGGCTGTGGAGAGGCGGCTCAAAATGCCAAATCAGAGATTACAAAGAAATCGGC 1920
2149 GCACTTGGCTGTGGAGAGGCGGCTCAAAATGCCAAATCAGAGATTACAAAGAAATCGGC 2208
1921 GATGTCATGCTGAAGAACATTCAGGSCATGAAGCACTTGGCGGCTCACCTGTGAAGCAC 1980
2209 GATGTCATGCTGAAGAACATTCAGGSCATGAAGCACTTGGCGGCTCACCTGTGAAGCAC 2268
1981 AGCGAGGCTTTGAGAGGCGCTTGAAGATGAATCAAGAGCTCCCGGCGCTGAGAACTTC 2040
2269 AGCGAGGCTTTGAGAGGCGCTTGAAGATGAATCAAGAGCTCCCGGCGCTGAGAACTTC 2328
2041 TGCAGAGCTTTGAGCTCAGAGAGGTGTTACCTACAGGCTCAACCTTCCCTCGTGGCG 2100
2329 TGCAGAGCTTTGAGCTCAGAGAGGTGTTACCTACAGGCTCAACCTTCCCTCGTGGCG 2388
2101 CCACTGACACGGGCTCATGCACTACAGAGGCTCTGGAAGCGCTGTGCAACCAACCCG 2160
2389 CCACTGACACGGGCTCATGCACTACAGAGGCTCTGGAAGCGCTGTGCAACCAACCCG 2448
2161 CCGAGCCACGCGGACTTCAAGGACCTGCCAGGCGCTTTGGCAGAGATCAAGAGATGGTG 2220
2449 CCGAGCCACGCGGACTTCAAGGACCTGCCAGGCGCTTTGGCAGAGATCAAGAGATGGTG 2508
2221 GCAGAGCTCCAGGATGATGATCAAGATGAGAAATTTCCAGAGCTGACAGAACTCAAG 2280
2509 GCAGAGCTCCAGGATGATGATCAAGATGAGAAATTTCCAGAGCTGACAGAACTCAAG 2568
2281 AAAGATTTGATTTGGATTTGACAAATCTTGTGTTCCGGGAAGGAGATTCTCGTGGGC 2340
2569 AAAGATTTGATTTGGATTTGACAAATCTTGTGTTCCGGGAAGGAGATTCTCGTGGGC 2628
2341 AGCTTCAGCAAGCTTCCGGGGAAGGGGCTCCAGCAGGCGCATTTTCTGTTCAGACAC 2400
2629 AGCTTCAGCAAGCTTCCGGGGAAGGGGCTCCAGCAGGCGCATTTTCTGTTCAGACAC 2688
2401 GTCTGCTATACAGAGCGGGGCTGAGCGGCTTCAATTCAGTTTAAAGTCCAAGGAG 2460
2689 GTCTGCTATACAGAGCGGGGCTGAGCGGCTTCAATTCAGTTTAAAGTCCAAGGAG 2748
2461 CTCCGCTCTATGAGCATGATTAAGAGAGCGAAGCAGATGGGGGTGCCCACTGC 2520
2749 CTCCGCTCTATGAGCATGATTAAGAGAGCGAAGCAGATGGGGGTGCCCACTGC 2808
2521 CTGACCTCTCCGGGACAGCGGCACTTCATCTGTGGCCGCACTTCTCGGTCCGAGATG 2580
2809 CTGACCTCTCCGGGACAGCGGCACTTCATCTGTGGCCGCACTTCTCGGTCCGAGATG 2868
2581 GAGAGTGGGTTGAGAGCATCAATGAGCTGCAATTTGACTTGGCGGAGAGAGAGAGAGCC 2640
2869 GAGAGTGGGTTGAGAGCATCAATGAGCTGCAATTTGACTTGGCGGAGAGAGAGAGAGCC 2928
2641 GCCCTGAGTTCTTGGCCAGAGGCCCTGTAACAAGATCCCTGATGAAGCCACCGCG 2700

Db 522 GCTCAAGCAGGTTGACGTGTAATGACACAGCGACGCTCTCTTGATTTCAACATTGTG 581
QY 541 CAATCTGAGATGGGGATTTTGAATGAAGCTTGAGACAGAGAGACTTACCAAAAATAA 600
Db 582 CAATCTGAGATGGGGATTTTGAATGAAGCTTGAGACAGAGAGACTTACCAAAAATAA 641
QY 601 TACATTAACCTCAGACAGACCACTAGAGACAATAATCGTGGAAATTTCCACATAACCAATT 660
Db 642 TACATTAACCTCAGAGAACCACTAGAGACAATAATCGTGGAAATTTCCACATAACCAATT 701
QY 661 GGAACAAACACAGACAGATCAGATTTTCAAGCTCTTAAGATTGCCGTGCGCTAGAGATG 720
Db 702 GGACAAACACAGACAGATCAGATTTTCAAGCTCTTAAGATTGCCGTGCGCTAGAGATG 761
QY 721 TATGGAATCCGGTTGACCCGCGCAGAGACAGGGAAGGACAGAGATTCATCTGCGCGTT 780
Db 762 TATGGAATCCGGTTGACCCGCGCAGAGACAGGGAAGGACAGAGATTCATCTGCGCGTT 821
QY 781 GCCAACACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCG 840
Db 822 GCCAACACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCG 881
QY 841 AAGGTGCGGAAGTGAAGCTTCAAGAGAAAGCGTTTCTCATCAAGCTCCGCGCAGATGCC 900
Db 882 AAGGTGCGGAAGTGAAGCTTCAAGAGAAAGCGTTTCTCATCAAGCTCCGCGCAGATGCC 941
QY 901 AATPAGTGGTACAGGATACCTTGGAATTCCTGATGGCCAGTCGGGATTTCTGCAAGTCC 960
Db 942 AATPAGTGGTACAGGATACCTTGGAATTCCTGATGGCCAGTCGGGATTTCTGCAAGTCC 1001
QY 961 TTCTGGAAATCTGTGTGTAACATCATGCTTCTTTAGAATTGTAAGAGCCCAACCA 1020
Db 1002 TTCTGGAAATCTGTGTGTAACATCATGCTTCTTTAGAATTGTAAGAGCCCAACCA 1061
QY 1021 AAGCCCAAGCCGCTCTTTAGCCGGGGGTCACTATTTGCGTTCAGTGTGGAATCGAG 1080
Db 1062 AAGCCCAAGCCGCTCTTTAGCCGGGGGTCACTATTTGCGTTCAGTGTGGAATCGAG 1121
QY 1081 AAGCAGGTTCTGACTATGTTAAAGAAAGAGACATAAGAAAGTGCAGTTTGAAGAGAG 1140
Db 1122 AAGCAGGTTCTGACTATGTTAAAGAAAGAGACATAAGAAAGTGCAGTTTGAAGAGAG 1181
QY 1141 CACAGCAGATTCATTCTATCCGAGACCTTGCTTCAAGCTTACAGAACTGAATTCGAA 1200
Db 1182 CACAGCAGATTCATTCTATCCGAGACCTTGCTTCAAGCTTACAGAACTGAATTCGAA 1241
QY 1201 GTGCTGAGACAGTCTTACAGAGACACAGCTTACATTTGGAGAAAGTCCGCAATCTCA 1260
Db 1242 GTGCTGAGACAGTCTTACAGAGACACAGCTTACATTTGGAGAAAGTCCGCAATCTCA 1301
QY 1261 GGGGGGCAAGGCTGCGCGGAGAAAGAAAGCAAGATTTTCGCGCGGGAGCCGGGGGTG 1320
Db 1302 GGGGGGCAAGGCTGCGCGGAGAAAGAAAGCAAGATTTTCGCGCGGGAGCCGGGGGTG 1361
QY 1321 CACCCGAGGCTTGCGCGGAGAAAGCCCGCGGGTAAACAGACAGGCGGACGAGCGGCG 1380
Db 1362 CACCCGAGGCTTGCGCGGAGAAAGCCCGCGGGTAAACAGAGAGCGGACGAGCGGCG 1421
QY 1381 TGGGCGCCCAAGAGAAAGAGAGAGTCTTTAAGATTAAGACCCAGAGAGTAAACTT 1440
Db 1422 TGGGCGCCCAAGAGAAAGAGAGAGTCTTTAAGATTAAGACCCAGAGAGTAAACTT 1481
QY 1441 CAGCCCCCGCAGCAAGACAGAGCTCCCTGACATGGCAGTCTCAACTTTCCGAGCTGTCT 1500
Db 1482 CAGCCCCCGCAGCAAGACAGAGCTCCCTGACATGGCAGTCTCAACTTTCCGAGCTGTCT 1541
QY 1501 GTGAATCTGCAAGGGGAGATGGCCCTTGCAAGCTGACCTTCTCCAACTGAGACCCC 1560
Db 1542 GTGAATCTGCAAGGGGAGATGGCCCTTGCAAGCTGACCTTCTCCAACTGAGACCCC 1601
QY 1561 GACACCAAGCAGGCTCTCTCTTTGATCAGCCGCTGCTGAATGACAGAGCTGCCCCGG 1620
Db 1602 GACACCAAGCAGGCTCTCTCTTTGATCAGCCGCTGCTGAATGACAGAGCTGCCCCGG 1661

QY 1621 ACGACGATGAGATGAGGGCCGAGAGAGAGATTTCCCACTGATTAAGCGTACTTCATA 1680
Db 1662 ACGACGATGAGATGAGGGCCGAGAGAGAGATTTCCCACTGATTAAGCGTACTTCATA 1721
QY 1681 GCTAAGAAAGTGTCTAACACCGAGCGAACATATCTGAAGAGATCTCGAAGTTATCACTTGG 1740
Db 1722 GCTAAGAAAGTGTCTAACACCGAGCGAACATATCTGAAGAGATCTCGAAGTTATCACTTGG 1781
QY 1741 TGGTTTCAGAGCAAGTGAAGCAAAAGAGACCGCATGCGGGAAGCACTGAAAAGTCTCATTA 1800
Db 1782 TGGTTTCAGAGCAAGTGAAGCAAAAGAGACCGCATGCGGGAAGCACTGAAAAGTCTCATTA 1841
QY 1801 TTCCGGAATTTTGAACCTTTTGACAAATTTCTATACATACTTCTCAAGGAAATTTGAGCAA 1860
Db 1842 TTCCGGAATTTTGAACCTTTTGACAAATTTCTATACATACTTCTCAAGGAAATTTGAGCAA 1901
QY 1861 CGACTTGCCCTGTGGAAAGCCGCTCAATGCCCCAATTCAGATTTACCAAGAAATCGGC 1920
Db 1902 CGACTTGCCCTGTGGAAAGCCGCTCAATGCCCCAATTCAGATTTACCAAGAAATCGGC 1961
QY 1921 GATGTCAATGCTGAAGAACTTTCAGGGCATGAGCACCTGGCGGCTCACTGTGGAAGCAC 1980
Db 1962 GATGTCAATGCTGAAGAACTTTCAGGGCATGAGCACCTGGCGGCTCACTGTGGAAGCAC 2021
QY 1981 AGCAGAGCCTTGAGAGCCCTGGAAGATGAAATCAAGAGCTCCGGCGGCTGGAAGAACTTC 2040
Db 2022 AGCAGAGCCTTGAGAGCCCTGGAAGATGAAATCAAGAGCTCCGGCGGCTGGAAGAACTTC 2081
QY 2041 TGCAGAGACTTTGAGCTGCAAGAGGTGTGTTACTTACCGCTCAACACTTCTCTGCGG 2100
Db 2082 TGCAGAGACTTTGAGCTGCAAGAGGTGTGTTACTTACCGCTCAACACTTCTCTGCGG 2141
QY 2101 CCAGTCACCCGCTCATGCACTACAGAGGCTCTGAGCGGCTGGAACCAACCAACCCG 2160
Db 2142 CCAGTCACCCGCTCATGCACTACAGAGGCTCTGAGCGGCTGGAACCAACCAACCCG 2201
QY 2161 CCGAGCCACGCCGACTTGAGGAGCTGCGAGCGCGCTTTGCGCAGAGATCAGGAGATGAGT 2220
Db 2202 CCGAGCCACGCCGACTTGAGGAGCTGCGAGCGCGCTTTGCGCAGAGATCAGGAGATGAGT 2261
QY 2221 GCACAGCTTCACGATTCAGATGATCAAGATGAGAAATTTCCAGAGCTGCAGACCTCAAG 2280
Db 2262 GCACAGCTTCACGATTCAGATGATCAAGATGAGAAATTTCCAGAGCTGCAGACCTCAAG 2321
QY 2281 AAAGATTTGATTTGATTTGAATGAACAATCTGTGTTCCGGGAAGGGAAGTCAATCCGTGCGG 2340
Db 2322 AAAGATTTGATTTGATTTGAATGAACAATCTGTGTTCCGGGAAGGGAAGTCAATCCGTGCGG 2381
QY 2341 AGCCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGACGCGCATGTTCTCTGTTCAACGAC 2400
Db 2382 AGCCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGACGCGCATGTTCTCTGTTCAACGAC 2441
QY 2401 GTCTGTCTATACAGAGCCGGGGGCTGACGCGCTTCAATCAATTTAAAGTCCAAGGGCAG 2460
Db 2442 GTCTGTCTATACAGAGCCGGGGGCTGACGCGCTTCAATCAATTTAAAGTCCAAGGGCAG 2501
QY 2461 CTCCCGCTCTATGGAATGATGAGAGAGCGAAGACGAGTGGGGGGTGGCCCACTGC 2520
Db 2502 CTCCCGCTCTATGGAATGATGAGAGAGCGAAGACGAGTGGGGGGTGGCCCACTGC 2561
QY 2521 CTGACCTTCGGGGGCAAGCGGAGTCAATCATCTGAGCGCGCAATTCCTGCTCCGAGATG 2580
Db 2562 CTGACCTTCGGGGGCAAGCGGAGTCAATCATCTGAGCGCGCAATTCCTGCTCCGAGATG 2621
QY 2581 GAGAAAGTGGTTGAGAGCAATCCAGATGCGCATTTGACTTGCGGGAAGAGAGCAGACGCC 2640
Db 2622 GAGAAAGTGGTTGAGAGCAATCCAGATGCGCATTTGACTTGCGGGAAGAGAGCAGACGCC 2681
QY 2641 GCCCCTGAGTTCTTGGCGAGACGCCCTGCAACAAGTCCCTGATGAAGCAGACCGG 2700
Db 2682 GCCCCTGAGTTCTTGGCGAGACGCCCTGCAACAAGTCCCTGATGAAGCAGACCGG 2741

QY 2701 GCTGACGAGAGTCAAGATGACCTGAGCGCTCGGCACTGCTGAGCGCCAGGCC 2760
 DB 2742 GCTGACGAGAGTCAAGATGACCTGAGCGCTCGGCACTGCTGAGCGCCAGGCC 2801
 QY 2761 CCGCACCAGGCGCAACAATGATGACCTGAGCGCGCAACAACAAGCTCTCCATG 2820
 DB 2802 CCGCACCAGGCGCAACAATGATGACCTGAGCGCGCAACAACAAGCTCTCCATG 2861
 QY 2821 GTCGACTTCAAGCTCCGATGAGGAATCACTTGTCTGAGAACTCTGAGAGAAATTCAA 2880
 DB 2862 GTGGACTTCAAGCTCCGATGAGGAATCACTTGTCTGAGAACTCTGAGAGAAATTCAA 2921
 QY 2881 AACAGCAAGGAGGAGGAGAGAGCTGAGGAGTGTTCACAACTTCTGAGCTCTTCTAC 2940
 DB 2922 AACAGCAAGGAGGAGGAGAGAGCTGAGGAGTGTTCACAACTTCTGAGCTCTTCTAC 2981
 QY 2941 AATACACACAGGACAAATCACTCCCTTGCAGCTCTCTGCTGAGCTCTCACC 3000
 DB 2982 AATACACACAGGACAAATCACTCCCTTGCAGCTCTCTGCTGAGCTCTCACC 3041
 QY 3001 ATCCCTCTGAGCTCCGAGAAATCCAGAAAGACTGATGATCAAGCTCTCAAGTCC 3060
 DB 3042 ATCCCTCTGAGCTCCGAGAAATCCAGAAAGACTGATGATCAAGCTCTCAAGTCC 3101
 QY 3061 CAGCTCTACTCTTCAAGGCGGAGAGAGTACAGCTTCAAGAGTGTGAGTGAAGTATC 3120
 DB 3102 CAGCTCTACTCTTCAAGGCGGAGAGAGTACAGCTTCAAGAGTGTGAGTGAAGTATC 3161
 QY 3121 CCGAGTCCGACCAAGCTCTGCTGCGCAAGCCCAAGTGTGAGCAACAAGTCTCTGAG 3180
 DB 3162 CCGAGTCCGACCAAGCTCTGCTGCGCAAGCCCAAGTGTGAGCAACAAGTCTCTGAG 3221
 QY 3181 TATTGATGCGCGGAGCAACATCTGTTTCCGAGTGTGAGTGTGAGTGTGAGTGTGAG 3240
 DB 3222 TATTGATGCGCGGAGCAACATCTGTTTCCGAGTGTGAGTGTGAGTGTGAGTGTGAG 3281
 QY 3241 CTCTCTATTATTAAGAGCTGTGTAATTAACCTGTCTGAAATCAAAATCATGAGCTT 3300
 DB 3282 CTCTCTATTATTAAGAGCTGTGTAATTAACCTGTCTGAAATCAAAATCATGAGCTT 3341
 QY 3301 CCGAGAGCTCTCTCTGTCACAGCGCGCTTTTAAACCCGAGCTCTCAGCGTTGAA 3360
 DB 3342 CCGAGAGCTCTCTCTGTCACAGCGCGCTTTTAAACCCGAGCTCTCAGCGTTGAA 3401
 QY 3361 TGAACAGCGCTCCCACTCTGAGTCTGAGTCCGCTGAGGCGCTGTTCTTTAGCTAGT 3420
 DB 3402 TGAACAGCGCTCCCACTCTGAGTCTGAGTCCGCTGAGGCGCTGTTCTTTAGCTAGT 3461
 QY 3421 CCAAGTATTAACATGTCATT 3442
 DB 3462 CCAAGTATTAACATGTCATT 3483

RESULT 5
 LOCUS BCO41595 3546 bp mRNA linear pri 08-MAR-2005
 DEFINITION Homo sapiens PRRM, RHOGEF (ARRGEF) and plectectrin domain protein 1
 (chondrocyte-derived), transcript variant 1, mRNA (cDNA clone
 MGC:52223 IMAGE:5443230), complete cds.
 ACCESSION BCO41595
 VERSION BCO41595.1 GI:27370806
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 1 (bases 1 to 3546)
 Strauberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
 Collinge FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg
 B, Buelow KH, Schaefer CE, Bhat NK, Hopkins RF, Jordan H, Moore T,
 Max SI, Wang J, Helein F, Dlatchenko L, Marusina K, Farmer AA, Rubin
 GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL,

JOURNAL
 PUBLISHED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 2 (bases 1 to 3546)
 Director MGC Project.
 Submitted (20-DEC-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akheri N., Ayele K., Beckstrom-Sternberg S.M., Benjamin B.,
 Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S.,
 Dietrich N.L., Granite S., Guan X., Gupta J., Haghighi P.,
 Hansen N., Ho S.-L., Karlins E., Kwong P., Latic P., Legaspi R.,
 Maduro Q.L., Mastello C., Maskeri B., Mastrian S.D., McCloskey J.C.,
 McDowell J., Pearson R., Stantrop S., Thomas P.J., Touchman J.W.,
 Tsurgeon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.,
 Young A., Zhang L.-H. and Green E.D.

REMARK
 COMMENT
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRIL Plate: 44 Row: i Column: 24.

FEATURES
 source
 Location/Qualifiers
 1..3546

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:52223 IMAGE:5443230"
 /cfeature_type="Brain, astrocytoma, grade IV"
 /clone_id="N1H_MGC_98"
 /lab_host="DH10B"
 /note="Vector: pOTB7"
 1..3546
 /gene="PRRP1"
 /note="synonyms: CDEP, MGC87400"
 /db_xref="GeneID:10160"
 /db_xref="MIM:602654"
 266..3403
 /gene="PRRP1"
 /codon_start=1
 /product="PRRM, RHOGEF, and plectectrin domain protein 1,
 isoform 1"
 /protein_id="AAH41595.1"
 /db_xref="GI:49898405"
 /db_xref="GeneID:10160"
 /db_xref="MIM:602654"
 /translation="MGIKQPTTPGRLCAPENSIGSTLERQKPPPTPSGLVSIK
 QMLDITQENAEVQPRAPGVLLDAVCNHNLTAVGDFGLFPPDHKITYLWLDLKFIV
 KQIRREKHVVAVFPPDPHTOLOEITRYLFLAQVQDAQGLTCDNSAALLI
 SHIVSEIDPFDAEDREHLAKNKYIPQDALEDKIVFHHNHIGOTPAESDFOLLEI

CDS
 1..3546
 /gene="PRRP1"
 /note="synonyms: CDEP, MGC87400"
 /db_xref="GeneID:10160"
 /db_xref="MIM:602654"
 266..3403
 /gene="PRRP1"
 /codon_start=1
 /product="PRRM, RHOGEF, and plectectrin domain protein 1,
 isoform 1"
 /protein_id="AAH41595.1"
 /db_xref="GI:49898405"
 /db_xref="GeneID:10160"
 /db_xref="MIM:602654"
 /translation="MGIKQPTTPGRLCAPENSIGSTLERQKPPPTPSGLVSIK
 QMLDITQENAEVQPRAPGVLLDAVCNHNLTAVGDFGLFPPDHKITYLWLDLKFIV
 KQIRREKHVVAVFPPDPHTOLOEITRYLFLAQVQDAQGLTCDNSAALLI
 SHIVSEIDPFDAEDREHLAKNKYIPQDALEDKIVFHHNHIGOTPAESDFOLLEI

Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Dr. Stefan Hansson
 cDNA Library Preparation: Michael Brownstein / Ted Udén
 Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mdickpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAX Plate: 168 Row: 1 Column: 8
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 48528036.

FEATURES

source

1..4932
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:87400 IMAGE:30342239"
 /issue_type="Placenta, normal"
 /clone_1b="NIH_MGC_147"
 /lab_host="DH10B"
 /note="Vector: pBluescriptR"

gene

CDS

1..4932
 /gene="FARP1"
 /note="synonyms: CDEP, MGC87400, PLEKHC2"
 /db_xref="GeneID:10160"
 /db_xref="MIM:602654"
 240..3470
 /gene="FARP1"
 /codon_start=1
 /product="FARP1 protein"
 /protein_id="AAH71592.1"
 /db_xref="GI:47940452"
 /db_xref="GeneID:10160"
 /db_xref="MIM:602654"
 /translation="MGEIQRPTPGSRILCAPENSIGSTIEBROKPPPTPSGLVSIKI
 QMLDPTQEAFEVQPAHGVKLVLDVACNHLNLTVEGDFGLPEPDHKITVWLKLEIV
 KQIRPKHVVVVKVFPFDDHLOQLBETRLVFLQVQDLAQGLTNDTSALLI
 SHIYSEIDFDEALDRHLAKNKYIPQDADLEDKIVENHNHIGQTPAESFOLEI
 ARLEWYGLRIHAPAKDRBGKTLNLANVANGILVQGFITLNFNNAKVKSIFKKRF
 LTKAPDANSAYODTLEPLMASRDFECKSPWKI CVENHAFPLFBEPKPKPVLRSRG
 SSFPRSGRTQKQVLDVKEGKHQVQFEKSHKISHIRSLAQPTIEINSEVLEQSQS
 TSLTGEASPEGGSCRRGKPKVSAGPSSHPSAPRRSPAGNKQADGASAPTEE
 BEEVYDRTOQSPQPOPSTGLTGSPLSLSVNSQGVAPAVNTLSPNLSPTKQ
 ASPLISPLNDACRTDEDEGRKRKPTDRAVFAKESVTERVYLKDLIVITSMF
 OSTVSKEDMPALSKSLIPNPEPLHKFTPTNLKEIQRALMEGRSNQKIDYORIG
 DVMLEKIOGMKHLAHLMKVSRLEALENGITKSSRLREHRCDFPELOKVCYLLPMLFL
 LRLRLHMYKQVLERLCKHHPSSHADPDCRAALAEITBEMVQLHGTMTKBNOKL
 HELKDLIGIDNLVVGPRPSSLSMKRTPHLDGAKRI PCAPERPLLVKEFTRLSLS
 KLSGKLOQRMFLFNDVLLYTSRGITLTAQNPQVHQLLPCAYMTIEESPDENGVPCL
 TLRGROSLIIVAASSSEMEKWEVDIQMIDLAELKSSLPAPFLTSPDNGSPBAT
 AADQSEDDLSASRTSLERQAPHRGNTVSHMNTVSMTDFSIAYENQSGMLIK
 KKNNGMKMLAVFTNPLTFYKSHODNHPLASLPLIGSLTISESNIQKDYVFK
 LHFKSHVYFRABSEYTFPERMVEVIRSAITSSASRPVLSHKESLYV"

ORIGIN

Query Match 96.2%; Score 3310.2; DB 5; Length 4932;
 Best Local Similarity 97.3%; Pred. No. 0;
 Matches 3415; Conservative 0; Mismatches 3; Indels 93; Gaps 1;

QY 25 GATATTCTTAAAGCCGCTTCATCATGGAGAAATGAGACAGGCCGACCCAGATCA 84
 DB 216 GATATTCTTAAAGCCGCTTCATCATGGAGAAATGAGACAGGCCGACCCAGATCA 275
 QY 85 CGACTGGGGGCCCCGGAATAATCGGGGATCAGTACCTTGGAACTGGAGAGAAAGCCGCC 144
 DB 276 CGACTGGGGGCCCCGGAATAATCGGGGATCAGTACCTTGGAACTGGAGAGAAAGCCGCC 335

QY 145 CCACACCTTCAGAGAAAATCTGTGTCATCAAAATCCAGATCTGATGACACCCAGAG 204
 DB 336 CCAACACCTTCAGAGAAAATCTGTGTCATCAAAATCCAGATCTGATGACACCCAGAG 395
 QY 205 GATTTTGAAGTTCCCAAAAGAGCTCTGGGAAGGTGCTGGGATTTGGCAACAC 264
 DB 366 GATTTGAAGTTCCCAAAAGAGCTCTGGGAAGGTGCTGGGATTTGGCAACAC 455
 QY 265 CTCAACCTCGTGGAGGTGACTATTTTGGCCCTGAGATTCTGATCACAAAAAGATCAG 324
 DB 456 CTCACCTCGTGGAGGTGACTATTTTGGCCCTGAGATTCTGATCACAAAAAGATCAG 515
 QY 335 GTTGGCTGGATCTCTTAAACCCATTGTGAACAGATTGAAAGGCCAAAGCAGTTGTT 384
 DB 516 GTTGGCTGGATCTCTTAAACCCATTGTGAACAGATTGAAAGGCCAAAGCAGTTGTT 575
 QY 385 GTTAAGTTTGTGTGAATTTCTTCCGCTGACCAACAACTCCAGAAAGATCACA 444
 DB 576 GTTAAGTTTGTGTGAATTTCTTCCGCTGACCAACAACTCCAGAAAGATCACA 635
 QY 445 AGTACCTGTTGCGCTGAGGTGAAGCAGAGACTTGGCTCAAGGAGGTTGACGTGTAAT 504
 DB 636 AGTACCTGTTGCGCTGAGGTGAAGCAGAGACTTGGCTCAAGGAGGTTGACGTGTAAT 695
 QY 505 GACACACGAGCAGCTCTCTTGAATTTACACATTTGTGCAATCTGAGATTGGGATTTGAT 564
 DB 696 GACACACGAGCAGCTCTCTTGAATTTACACATTTGTGCAATCTGAGATTGGGATTTGAT 755
 QY 565 GAAGCTTTGAGACAGAGACACTTACCAAAAAATAATCATCTCAGCAAGACGACAT 624
 DB 756 GAAGCTTTGAGACAGAGACACTTACCAAAAAATAATCATCTCAGCAAGACGACAT 815
 QY 625 GAGACAAAATCGTGAATTTTACCATTAACCATTTGACAAACACAGCAGAAATCAGAT 684
 DB 816 GAGACAAAATCGTGAATTTTACCATTAACCATTTGACAAACACAGCAGAAATCAGAT 875
 QY 685 TTCCAGCTCTTGAAGATTTCCCGCTGAGATGATGATCCGGTTGACACCCGCC 744
 DB 876 TTCCAGCTCTTGAAGATTTCCCGCTGAGATGATGATCCGGTTGACACCCGCC 935
 QY 745 AAGACAGGGAAGGACCAAGATTCATCTGCGCTTGCACAACGGGAATTCATAGTTT 804
 DB 936 AAGACAGGGAAGGACCAAGATTCATCTGCGCTTGCACAACGGGAATTCATAGTTT 995
 QY 805 CAGGTTTCACTAATGATCAATCTCTTCAATGGGCAAGGTGCGGAGCTTGACTTCAAG 864
 DB 996 CAGGTTTCACTAATGATCAATCTCTTCAATGGGCAAGGTGCGGAGCTTGACTTCAAG 1055
 QY 865 AGGAAGCGTTTCTCATCAAGCTCCGGCCAGATGCCAATAGTGCCTACAGATACCTTG 924
 DB 1056 AGGAAGCGTTTCTCATCAAGCTCCGGCCAGATGCCAATAGTGCCTACAGATACCTTG 1115
 QY 925 GAATTCCTGATGCGACAGTCGGGATTTTCGCAAGTCTCTGAAAAATCTGTGTGAACAT 984
 DB 1116 GAATTCCTGATGCGACAGTCGGGATTTTCGCAAGTCTCTGAAAAATCTGTGTGAACAT 1175
 QY 985 CATGCTCTTTTGAAGCTTTTGAAGGCCCAACCAAGGCCAGCCCTCTCTTATAC 1044
 DB 1176 CATGCTCTTTTGAAGCTTTTGAAGGCCCAACCAAGGCCAGCCCTCTCTTATAC 1235
 QY 1045 CGGGGGTCAATCATTTTCGTTCAAGTGTGGACTCAAGACAGGTTCTGACTATGTTAA 1104
 DB 1236 CGGGGGTCAATCATTTTCGTTCAAGTGTGGACTCAAGACAGGTTCTGACTATGTTAA 1295
 QY 1105 GAAGAGGACATTAAGAGGTGAGTTTGAAGAGGACAGAAAGATTGATCTATCCGG 1164
 DB 1296 GAAGAGGACATTAAGAGGTGAGTTTGAAGAGGACAGAAAGATTGATCTATCCGG 1355
 QY 1165 AGCTTGTCTTCAACAGCTTACAGAACTGAATTCGAAAGTCTGAGAGCTCTCAGCAGAG 1224
 DB 1356 AGCTTGTCTTCAACAGCTTACAGAACTGAATTCGAAAGTCTGAGAGCTCTCAGCAGAG 1415
 QY 1225 ACCAGCTTACATTTTGAAGAGGTGCCAATCTCCAGGGGGCCAGAGCTCCGGCGAGGA 1284

Db 1416 ACACAGCTTACATTTGGAGAAAGGTGCGAATCTCCAGGGGGCCAGAGCTGCCGGCAGAGA 1475
Qy 1285 AAGGAACCGAAGGTTTCCCGCGGGAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGA 1344
Db 1476 AAGGAACCGAAGGTTTCCCGCGGGAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGA 1535
Qy 1345 AGCCCGCGGGGTAAACAAGCAGAGGGGAGCCGAGCGCGCTCGGGCGCCACCGAGGAAGAGAG 1404
Db 1536 AGCCCGCGGGGTAAACAAGCAGAGGGGAGCCGAGCGCGCTCGGGCGCCACCGAGGAAGAGAG 1595
Qy 1405 GAGGTGTTAAGGATAGAACCCAGAGAGTAAACTTCAGCCCCCGAGCCCAACAGAGC 1464
Db 1596 GAGGTGTTAAGGATAGAACCCAGAGAGTAAACTTCAGCCCCCGAGCCCAACAGAGC 1655
Qy 1465 TCCCTGACTGGCAGTCTCTCACTTTCCGAGCTGTCTGTAACTCGCAGGGGGAGTGGCC 1524
Db 1656 TCCCTGACTGGCAGTCTCTCACTTTCCGAGCTGTCTGTAACTCGCAGGGGGAGTGGCC 1715
Qy 1525 CCGGCAAGTACCTTGTCTCCCACTGAGCCCGAGCCGAGCAGAGAGGCTCTCCCTTG 1584
Db 1716 CCGGCAAGTACCTTGTCTCCCACTGAGCCCGAGCAGAGAGGCTCTCCCTTG 1775
Qy 1585 ATCAGCCCGCTGCTGATGACAGAGCCTGCCCGAGCGAGCGATGAGATGAGGCGCG 1644
Db 1776 ATCAGCCCGCTGCTGATGACAGAGCCTGCCCGAGCGAGCGATGAGATGAGGCGCG 1835
Qy 1645 AAGAGAGATTTCCCACTGATTAAGGCTTCTATAGCTAAGAGAGTGTCTACACCGAG 1704
Db 1836 AAGAGAGATTTCCCACTGATTAAGGCTTCTATAGCTAAGAGAGTGTCTACACCGAG 1895
Qy 1705 CGAACAATCTGAGAGATCTCGAAGTTATCACTTGTGTTTGAAGCAGTGAAGAA 1764
Db 1896 CGAACAATCTGAGAGATCTCGAAGTTATCACTTGTGTTTGAAGCAGTGAAGAA 1955
Qy 1765 GAGAGCGCATGCGGAGAGCACTGAAAGTCTCATATTTCCGAAATTTTGAACCTTTGAC 1824
Db 1956 GAGAGCGCATGCGGAGAGCACTGAAAGTCTCATATTTCCGAAATTTTGAACCTTTGAC 2015
Qy 1825 AAATTTCTATCTAATTTTCTCAAGAAATTGAGCAACGACTTCCCTGCGGAGGCGCG 1884
Db 2016 AAATTTCTATCTAATTTTCTCAAGAAATTGAGCAACGACTTCCCTGCGGAGGCGCG 2075
Qy 1885 TCAAATGCCCAATCAAGATTACCAAGAAATGGGAGATGTCATGCTGAAGAACATTGAG 1944
Db 2076 TCAAATGCCCAATCAAGATTACCAAGAAATGGGAGATGTCATGCTGAAGAACATTGAG 2135
Qy 1945 GGCATGAAGCACTGCGGCTCACTGTGGAACAGAGCGAGGCTTTGAGGCGCTGAGAG 2004
Db 2136 GGCATGAAGCACTGCGGCTCACTGTGGAAGTACAGAGAGGCTTTGAGGCGCTGAGAG 2195
Qy 2005 AATGGAATCAAGAGCTCCCGCGGCTGAGAACTTCTGAGAGACTTTGAGCTGCAAGAG 2064
Db 2196 AATGGAATCAAGAGCTCCCGCGGCTGAGAACTTCTGAGAGACTTTGAGCTGCAAGAG 2255
Qy 2065 GTGTGTACTTACCGGCTCAACACTTCTCTCGGCGGCACTGACCGGCTCATGCACTAC 2124
Db 2256 GTGTGTACTTACCGGCTCAACACTTCTCTCGGCGGCACTGACCGGCTCATGCACTAC 2315
Qy 2125 AAGCAGAGCTCTGAGAGGCTGTGCAAAACACACCCCGAGCCAGCCGCACTTCAAGAGAC 2184
Db 2316 AAGCAGAGCTCTGAGAGGCTGTGCAAAACACACCCCGAGCCAGCCGCACTTCAAGAGAC 2375
Qy 2185 TGCAGAGCGCTTTGGCAGAGATCACGAGAGTGTGGCACAAGTCTCACCGTACGATGATC 2244
Db 2376 TGCAGAGCGCTTTGGCAGAGATCACGAGAGTGTGGCACAAGTCTCACCGTACGATGATC 2435
Qy 2245 AAGATGAGAAATTTCCAGAACTGCAAGAACTCAAGAAAGATTGATGATGCAATTCAGAA 2304
Db 2436 AAGATGAGAAATTTCCAGAACTGCAAGAACTCAAGAAAGATTGATGATGCAATTCAGAA 2495
Qy 2305 CTTGTGTTCCGGAAG----- 2321

Db 2496 CTTGTGTTCCGGAAGGCTGTGCTCTCTTTTCCCTTATGAGGACTCCACATTAGAGCAA 2555
Qy 2322 -----GAGTTCATC 2331
Db 2556 GCCAGGCGCATCCCATGTGTCTCCAGAGAGAGGCCATTGTTATTAAGTAAGAGATTATC 2615
Qy 2332 CGTCTGGGAGGCTCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGAGCGCATGTTCTCTG 2391
Db 2616 CGTCTGGGAGGCTCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGAGCGCATGTTCTCTG 2675
Qy 2392 TTCAACAGCTCCTGCTATACAGAGCCGGGGGCTGACGGCTTCAATCAGTTTAAAGTC 2451
Db 2676 TTCAACAGCTCCTGCTATACAGAGCCGGGGGCTGACGGCTTCAATCAGTTTAAAGTC 2735
Qy 2452 CAGGGCAGCTCCCGCTCTATGAGCAATGAGAGAGCCGAAGAGAGTGGGGGTG 2511
Db 2736 CAGGGCAGCTCCCGCTCTATGAGCAATGAGAGAGCCGAAGAGAGTGGGGGTG 2795
Qy 2512 CCCCACTGCTGACCTCTCGGGGCGCAGCGGAGTCCATCATCTGTGCGCGCAGTTCGCG 2571
Db 2796 CCCCACTGCTGACCTCTCGGGGCGCAGCGGAGTCCATCATCTGTGCGCGCAGTTCGCG 2855
Qy 2572 TCCGAGATGAGAAAGTGGGTTGAGGATTCACAGATGGCCATTGACCTGGCGAGAAAGC 2631
Db 2856 TCCGAGATGAGAAAGTGGGTTGAGGATTCACAGATGGCCATTGACCTGGCGAGAAAGC 2915
Qy 2632 AGAGAGCCCGGCTCTGAGTTCCTGTGCGAGAGGCCCCCTGACAAACAAAGTCCCTGATGA 2691
Db 2916 AGAGAGCCCGGCTCTGAGTTCCTGTGCGAGAGGCCCCCTGACAAACAAAGTCCCTGATGA 2975
Qy 2692 GCCACGCGGCTGACCAAGAGTGAAGAGATGACTGAGCGCTCGGCAATCGCTGAGAG 2751
Db 2976 GCCACGCGGCTGACCAAGAGTGAAGAGATGACTGAGCGCTCGGCAATCGCTGAGAG 3035
Qy 2752 CGCAGAGCCCGGACCGCGGCAACACAAATGTCACAGTGTCTGGACCGCAACCAAGC 2811
Db 3036 CGCAGAGCCCGGACCGCGGCAACACAAATGTCACAGTGTCTGGACCGCAACCAAGC 3095
Qy 2812 GTTCTCAATGATGACTTTCAGCATTCGAGATGAGAAATCAAGTGTCTGGAACTGTGAGG 2871
Db 3096 GTTCTCAATGATGACTTTCAGCATTCGAGATGAGAAATCAAGTGTCTGGAACTGTGAGG 3155
Qy 2872 AAATTCAAAAACGACAGCGGCTGAGAGAGCTGTGGGTGTTCACAACTTCTGCTG 2931
Db 3156 AAATTCAAAAACGACAGCGGCTGAGAGAGCTGTGGGTGTTCACAACTTCTGCTG 3215
Qy 2932 TTCTTCTAACAAATCACACAGAGCAATCAATCCCTTGCAGAGCTGTGCTCGGCTAC 2991
Db 3216 TTCTTCTAACAAATCACACAGAGCAATCAATCCCTTGCAGAGCTGTGCTCGGCTAC 3275
Qy 2992 TCGCTCAACATCCCTCTGTGAGTCCGAGAACTCAAGAAAGCTAAGTTCAGAGCTGCAC 3051
Db 3276 TCGCTCAACATCCCTCTGTGAGTCCGAGAACTCAAGAAAGCTAAGTTCAGAGCTGCAC 3335
Qy 3052 TTCAAGTCCACGTTACTACTTCAAGGCGGAAAGGAGTACAGCTTCCGAAAGTGGATG 3111
Db 3336 TTCAAGTCCACGTTACTACTTCAAGGCGGAAAGGAGTACAGCTTCCGAAAGTGGATG 3395
Qy 3112 GAAGTATCCGAGTCCACAGAGCTGTGCTGCGACCCCAAGTGTGAGCCACAAAGAG 3171
Db 3396 GAAGTATCCGAGTCCACAGAGCTGTGCTGCGACCCCAAGTGTGAGTCAAAAGAG 3455
Qy 3172 TCTCTTGTGATTTGATGAGCGGACACACTGCTTCCGAGTGGCTGCTTCTGGAAGAC 3231
Db 3456 TCTCTTGTGATTTGATGAGCGGACACACTGCTTCCGAGTGGCTGCTTCTGGAAGAC 3515
Qy 3232 GTTTCCTTCTCTGTATTAATGAAGCTGTGTAATTAACACTCTGTAATAATCAAAA 3291
Db 3516 GTTTCCTTCTCTGTATTAATGAAGCTGTGTAATTAACACTCTGTAATAATCAAAA 3575
Qy 3292 ACATGCTTCCAGAGCTCTCTGTCTCCACAGCGGCTTTTAAACCCCGACCTCTCA 3351
Db 3576 ACATGCTTCCAGAGCTCTCTGTCTCCACAGCGGCTTTTAAACCCCGACCTCTCA 3635

QY 3352 GCGTTTGATGAACAGCGCTCCACCTCCAGTCTCGCATCCGCTGGGGGCGCGTGTCTT 3411
|||||
Db 3636 GCGTCTGATGATGAACCGCTCCACCTCCAGTCTCGCATCCGCTGGGGGCGCGTGTCTT 3695
|||||
QY 3412 TAGCTAGTCCAGTATTAACATTGTCAAT 3442
|||||
Db 3696 TAGCTAGTCCAGTATTAACATTGTCAAT 3726
|||||
RESULT 7
LOCUS BC030329 3902 bp mRNA linear ROD 21-OCT-2003
DEFINITION Mus musculus PERM, Rhogef (Arhgef) and pleckstrin domain protein 1
(chondrocyte-derived), mRNA (cDNA clone IMAGE:5376197), partial
cds
BC030329
BC030329.1 GI:20987935
Mus musculus (house mouse)
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 3902)
Strausberg, R.D., Collins, F.S., Wagner, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L.H., Derge, J.G.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marsina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stableton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Uedlin, T.B., Toshiyuki, S.,
Carinci, P., Prange, C., Raha, S.S., Loggiano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.D., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shevchenko, Y.,
Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, V., Smallus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3902)
Strausberg, R.
Direct Submission
Submitted (07-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nsl.nih.gov
Akher, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakeley, R.W., Bouffard, G.G., Breen, K., Binkley, C., Brooks, S.,
Dierich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Larc, P., Legardi, R.,
Maduro, O.L., Masiello, C., Maskert, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantip, B., Thomas, P.J., Touchman, J.W.,
Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAP Plate: 65 Row: m Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

FEATURES
source

Location/Qualifiers
1..3902
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5376197"
/tissue_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month
old mouse, Taken by biopsy."
/clone_lib="NCI_CGAP_Mam2"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
<1..3902
/gene="Parp1"
/note="synonym: Cdep"
/db_xref="geneid:223254"
/db_xref="MGI:246173"
<1..2510
/gene="Parp1"
/product="Parp1 protein"
/protein_id="AAH30329.1"
/db_xref="GI:20987936"
/db_xref="geneid:223254"
/db_xref="MGI:246173"
/translation="TRPEVARLLEMYGRLRPAKDEGKINLAVNTGILVFOGFTK
INAPWAVRKLSEPRKFLIKRPDNNSSVDTLEPLMAGDPCKSPFKITVEHNAF
RLPEPRKPEKPVLPSSGSSFRSGRQKQVLDVYKGGKRVQPERKSKTHRS
LVSGTANSEVTPVPSASLTREGEESFGQSCQAKETACVLELGGPSALP
KSPSSKADGTVVPPSEEEBEGKDGIRSNPQPPQPSGLTSGPHELSINS
OGGAPAVVTLSPNLSPNKQASPLISPLNDQACPRDDEEGRRKRPPTDRAVYA
KEVSTRTYVLLSEVIAVSPQTSKDSMEALSKLIFPVEPLKHTNPKLIE
ORLALMRNSAHVGDYRGIDVNLKNIQCKHLAAHLMKSSALELSTIKSRR
LEHFRDELRNVCVLPNTPLRLPLHLMRYHLBRCKHPPNHADPRCCRALA
ETTEVVALDGLTMIWENPQKHEIKQDILGIDNVTEREIRLGSLSKSGQIQ
RMFPLFNDVLVYTSKGLTASNOFKHQGLPLYGWTIESEBEGVPHCTLGRQSI
IIVASSREMEKMEIDIMALDLAKESGPTPELASSPDKNSPDEATPAQOESDD
LSASRTSLERQAPHRGNTVMVCHMRSTSVSDPSIAVENQSLNGLRPFKNSGMQ
KLWVFTNFCLEFFYKSHODSHPLALPLILGSLTTPSEENHKKQYVFLHKKSHYY
FRASEYTERMEVIRSATSSASRAHLISKESHLY"

CDS

misc_feature
12477932
2 (bases 1 to 3902)
Strausberg, R.
Direct Submission
Submitted (07-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nsl.nih.gov
Akher, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakeley, R.W., Bouffard, G.G., Breen, K., Binkley, C., Brooks, S.,
Dierich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Larc, P., Legardi, R.,
Maduro, O.L., Masiello, C., Maskert, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantip, B., Thomas, P.J., Touchman, J.W.,
Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

misc_feature

misc_feature
12477932
2 (bases 1 to 3902)
Strausberg, R.
Direct Submission
Submitted (07-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nsl.nih.gov
Akher, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakeley, R.W., Bouffard, G.G., Breen, K., Binkley, C., Brooks, S.,
Dierich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Larc, P., Legardi, R.,
Maduro, O.L., Masiello, C., Maskert, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantip, B., Thomas, P.J., Touchman, J.W.,
Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

misc_feature

misc_feature
12477932
2 (bases 1 to 3902)
Strausberg, R.
Direct Submission
Submitted (07-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nsl.nih.gov
Akher, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakeley, R.W., Bouffard, G.G., Breen, K., Binkley, C., Brooks, S.,
Dierich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Larc, P., Legardi, R.,
Maduro, O.L., Masiello, C., Maskert, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantip, B., Thomas, P.J., Touchman, J.W.,
Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

QY 2661 CAGCCCCCTGACAAACAGTCCCTGATGAAAGCCACCGGGCTGACACAGAGTCAGAGGA 2720
| | | | |
Db 1985 CAGCCCCCTGACAAACAGTCCCTGATGAAAGCCACCGGGCTGACACAGAGTCAGAGGA 2044
| | | | |
QY 2721 TGAACCTGAGCGCTCGCGACATCGCTGAGCGCAGGCCCGCACCGCGGCAACAAT 2780
| | | | |
Db 2045 TGAACCTGAGCGCTCGCGACATCGCTGAGCGCGCAGGCCCGCACCGCGGCAACAAT 2104
| | | | |
QY 2781 GGTGCACTGTGTGTCGACACCGCAACACAGCTCTTCATGTGTGATTCATGACATCGCAAT 2840
| | | | |
Db 2105 GGTGCACTGTGTGTCGACACCGCAACACAGCTCTTCATGTGTGATTCATGACATCGCAAT 2164
| | | | |
QY 2841 GGAGAACTGATTGTCTGAGAAACCTGCTGAGAAATTCAAAAACAGCAAGGGGAGAGAA 2900
| | | | |
Db 2165 GGAGAACTGATTGTCTGAGAACTGCTGAGAAATTCAAAAACAGCAATGCTGAGAGAA 2224
| | | | |
QY 2901 GCTGTGGGTGTGTTCACAACTTCGCTGTTCTTCAAAATCACACACAGACAAATCA 2960
| | | | |
Db 2225 GCTGTGGGTGTGTATTCACCAACTTCGCTGTTCTTCAAAATCACACACAGACAGTCA 2284
| | | | |
QY 2961 TCCCTTGTGACGCTGCTCTGTCTGCTACTACTGCTCAACATCCCTTGTGATCGAGAA 3020
| | | | |
Db 2285 CCCCCCTTGACAGCTGCTCTGTCTGCTACTACTGCTCAACATCCCTTGTGATCGAGAA 2344
| | | | |
QY 3021 CATCCAGAAAGACTAGTGTTCAGGCTGACCTTCAAGTCCACGCTACTACTTCAAGGC 3080
| | | | |
Db 2345 CATCCAGAAAGACTAGTGTTCAGGCTGACCTTCAAGTCCACGCTACTACTTCAAGGC 2404
| | | | |
QY 3081 GGAAACGAGTACACGCTTGAAAAGTGATGAAAGTATCCGAGTCCACACGCTCTGC 3140
| | | | |
Db 2405 TGAAGCGAATACACATTTGAAAAGTGATGAAAGTATCCGAAAGTCCACACGCTCAAG 2464
| | | | |
QY 3141 CTCGCAACCCCACTGTGTGAGCCACAAAGTCTTGTGTATTTATGTCGGGACCAACT 3200
| | | | |
Db 2465 CTCCTCGGCTCACATCTTAGTCAAAAGATCTCACTGTGACGGCTGACAGCC 2524
| | | | |
QY 3201 C 3201
| | | | |
Db 2525 C 2525
| | | | |

RESULT 8
LOCUS CQ724059 2937 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 9993 from Patent W002068579.
ACCESSION CQ724059
VERSION CQ724059.1 GI:42284916
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
1
Venter, C.J., Adams, M.C., Li, P.W., and Myers, E.W.
Kite, such as nucleic acid arrays, comprising a majority of
humaneons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 9993 06-SEP-2002;
PE Corporation (NY) (US)
Location/Qualifiers

JOURNAL
FEATURES
source
1..2937
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 53.8%; Score 1851.8; DB 2; Length 2937;
Best Local Similarity 79.9%; Pred. No. 0;
Matches 2513; Conservative 0; Mismatches 77; Indels 555; Gaps 6;
QY 220 CAAAGAGCTCTGGAAAGGTGCTGTGATGACATTGCAACCACTCAACTGTGAA 279

Db
| | | | |
QY 1 CAAAGAGCTCTGGAAAGGTGCTGTGATGACATTGCAACCACTCAACTGTGAA 60
| | | | |
QY 280 GGTGACTATTTGGCTGCTGAGTTTCTGATCAAAAAAGATACGGTGTGGCTGATCTC 339
| | | | |
Db 61 GGTGACTATTTGGCTGCTGAGTTTCTGATCAAAAAAGATACG----- 105
| | | | |
QY 340 CTAAAAACCATTTGTGAACAGATTGAAGGCCAAAGCAGTGTGTTAATTTGTTGGTG 399
| | | | |
Db 106 -----GCCAGTCCCGGTG 120
| | | | |
QY 400 AAATCTTTCGCTGACACACAACTCAAGAAATCTCAAGGTAACCTGTTTCGG 459
| | | | |
Db 121 ACCTCCTGCTCTTTGCTGACGCCAGCTCTGTGATCTTCGAGTACCTGTTCCGG 180
| | | | |
QY 460 CTGCAAGTGAACAGACCTTGGCTCAAGCAGTGTGACGTGAATGACACAGCGAGCT 519
| | | | |
Db 181 CTGCAAGTGAACAGACCTTGGCTCAAGCAGTGTGACGTGAATGACACAGCGAGCT 240
| | | | |
QY 520 CTCTGATTTCAACATTTGTGCAATCTGAGATTGGGATTTTGTGATGAAAGCTTGAAGA 579
| | | | |
Db 241 CTCTGATTTCAACATTTGTGCAATCTGAGATTGGGATTTTGTGATGAAAGCTTGAAGA 300
| | | | |
QY 580 GAGCACTTAGCAAAAAATTAATACATCTCAAGAAAGCAGCACTAGAGAACAAATCTG 639
| | | | |
Db 301 GAGCACTTAGCAAAAAATTAATACATCTCAAGAAAGCAGCACTAGAGAACAAATCTG 360
| | | | |
QY 640 GAATTTCCACCAATACCACTTGGACAAACACAGCAATCATGATTTCCAGCTCCCTAG 699
| | | | |
Db 361 GAATTTCCACCAATACCACTTGGACAAACACAGCAATCATGATTTCCAGCTCCCTAG 420
| | | | |
QY 700 ATTGCGCTGCTGAGATGTATGGAATCCGTTGCAACCGGCAAGGACAGGAGAGG 759
| | | | |
Db 421 ATTGCGCTGCTGAGATGTATGGAATCCGTTGCAACCGGCAAGGACAGGAGAGG 480
| | | | |
QY 760 ACGAAGATCAATCTGCGCTTGCCAAACAGGGAATCTAGTGTTCAGGTTCACTAAG 819
| | | | |
Db 481 ACGAAGATCAATCTGCGCTTGCCAAACAGGGAATCTAGTGTTCAGGTTCACTAAG 540
| | | | |
QY 820 ATCAATGCTTCAACTGCGGCAAGGTCGGAAGCTGAGCTTCAAGAGGAAGGCTTCTC 879
| | | | |
Db 541 ATCAATGCTTCAACTGCGGCAAGGTCGGAAGCTGAGCTTCAAGAGGAAGGCTTCTC 600
| | | | |
QY 880 ATCAAGCTCCGCGCAGATGCCAATAGTGGTACAGAGATACCTTGAATTCCTGATGCG 939
| | | | |
Db 601 ATCAAGCTCCGCGCAGATGCCAATAGTGGTACAGAGATACCTTGAATTCCTGATGCG 660
| | | | |
QY 940 AGTGGGATTTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATCATGCTTCTTAA 999
| | | | |
Db 661 AGTGGGATTTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATCATGCTTCTTAA 720
| | | | |
QY 1000 CTTTTTGAAGAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGGTTCATCATTT 1059
| | | | |
Db 721 CTTTTTGAAGAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGGTTCATCATTT 780
| | | | |
QY 1060 CGATTCAAGTGTGCAAGTCTGAAAGCAGTTCCTGACTATGTTAAAGAGAGACATAAG 1119
| | | | |
Db 781 CGATTCAAGTGTGCAAGTCTGAAAGCAGTTCCTGACTATGTTAAAGAGAGACATAAG 840
| | | | |
QY 1120 AAGGTGCAATTGAAAGAGACACAGCAAGATTCATTCTATCGAGCTTGTCTACAG 1179
| | | | |
Db 841 AAGGTGCAATTGAAAGAGACACAGCAAGATTCATTCTATCGAGCTTGTCTACAG 900
| | | | |
QY 1180 CCTACAGAACTGAATTCGAAGTGTGAGAGTCTCAGACAGACACAGCTTACATTT 1239
| | | | |
Db 901 CCTACAGAACTGAATTCGAAGTGTGAGAGTCTCAGACAGACACAGCTTACATTT 960
| | | | |
QY 1240 GGAGAAAGTGTCCGAATCTCCAGAGGCGCAGAGCTGCGCGAGAAAGAACCGAAGGTT 1299
| | | | |
Db 961 GGAGAAAGTGTCCGAATCTCCAGAGGCGCAGAGCTGCGCGAGAAAGAACCGAAGGTT 1020
| | | | |
QY 1300 TCCGCGGAGAGCCGGGCTGCAACCCAGCCCTGCGCGAGAGAAAGCCCGCGGTTAC 1359
| | | | |

Db 1021 TCCGCGGGAGCCGGGGTGCACCCGAGCCCTGCGCGAGAGAAAGCCCGGGGTAA 1080
 Qy 1360 AAGCAGGCGGACCGGAGCCCGCTCGGGCCCAACGGAAGAGAGAGGTCTTAAAGAT 1419
 Db 1081 AAGCAGGCGGACCGGAGCCCGCTCGGGCCCAACGGAAGAGAGAGGTCTTAAAGAT 1140
 Qy 1420 AGGAGCCAGAGATTAACCTCAGCCCGGAGCGAAGACA----- 1461
 Db 1141 AGGAGCCAGAGATTAACCTCAGCCCGGAGCGAAGACAAGGTCCAGATCCGGGCT 1200
 Qy 1462 ----- 1461
 Db 1201 GCCAGAGGACGAGTACTTCATTCTTGAAGTACGAGTACGAGTAAAGCAATAC 1260
 Qy 1462 ----- 1461
 Db 1261 GACCTTCTTAGAGAACAGGCTCCGATCCGGTCCCGATCCGGTCCCAACA 1320
 Qy 1462 -----GCTCCCTGACTGGGAGTCTCAACCTTCCGAGCTGTCTGAAC 1506
 Db 1321 AATGACAAATTATGAGGGCTCCCTGACTGGGAGTCTCAACCTTCCGAGCTGTCTGAAC 1380
 Qy 1507 TCGCAGGGGGAGTGGCCCTGCGCAAGTGAACCTTGTCTCCCAACCTGAGCCCGACCC 1566
 Db 1381 TCGCAGGGGGAGTGGCCCTGCGCAAGTGAACCTTGTCTCCCAACCTGAGCCCGACCC 1440
 Qy 1567 AAGCAGGCTCTTCCCTTGAATCAGCCCGCTGTGAATGAACAGGCTGTCCCGGAGGAC 1626
 Db 1441 AAGCAGGCTCTTCCCTTGAATCAGCCCGCTGTGAATGAACAGGCTGTCCCGGAGGAC 1500
 Qy 1627 GATGAGATGAGGGCCGAGAGAGATTCACCACTGATTAAGCGTACTTATAGCTAAG 1686
 Db 1501 GATGAGATGAGGGCCGAGAGAGATTCACCACTGATTAAGCGTACTTATAGCTAAG 1560
 Qy 1687 GAAGTGTCTACCAACGAGCGAACATATCTGAGAGATCTGAGATTAACCTTGTGTGTT 1746
 Db 1561 GAAGTGTCTACCAACGAGCGAACATATCTGAGAGATCTGAGATTAACCTTGTGTGTT 1620
 Qy 1747 CAGAGCAGGTGAGCAAAAGAGAGCGCCATGCCGGAAGCACTGAAGGTCTCATATTTCCG 1806
 Db 1621 CAGAGCAGGTGAGCAAAAGAGAGCGCCATGCCGGAAGCACTGAAGGTCTCATATTTCCG 1680
 Qy 1807 AATTTGGAACCTTTGCAAAATTTCACTAATTTTCTCAAGGAATTTGAGACGACTT 1866
 Db 1681 AATTTTGAACCTTTGCAAAATTTCACTAATTTTCTCAAGGAATTTGAGACGACTT 1740
 Qy 1867 GCCCTGTGGAAAGCCGCTCAATGCCCAATCAGAGATTAACCAAGAAATCCGCGATGTC 1926
 Db 1741 GCCCTGTGGATGAGCTGACA----- 1761
 Qy 1927 ATGCTGAAGAACATTTAGGGGATGAAGCACCTGCGGCTCACTGTGAAGCACAGCGAG 1986
 Db 1762 ----- 1761
 Qy 1987 GCCTTGAAGAGCCCTGAGAGATGGAATCAAGAGCTCCCGGCGCTGAGAACTTCTGACAG 2046
 Db 1762 -----GCTGCTAGGGGAAGACCCCTGTCTG 1786
 Qy 2047 GACTTTGAGCTGCAGAAAGGTGTGTACTACGCGCTCAACACTTCTCTCTCGGCGCACTG 2106
 Db 1787 CTTTCCGCTGAGACTGTGTATGTCTTTT----- 1817
 Qy 2107 CACCGGCTATGACTACAAAGAGGTCTGAGAGCGGCTGTGCAAAACACACCCGCGAGC 2166
 Db 1818 ----- 1817
 Qy 2167 CACGCGCACTTCAGGAGCTGCGAGCGCTTTGGAGAGATCAAGAGATGTGTGACAG 2226
 Db 1818 -----CTTACAGCCGCTTTGGAGAGATCAAGAGATGTGTGACAG 1860
 Qy 2227 CTCACAGGTACGATGATCAAGATGAGAAATTTCCAGAGCTGACGAACTCAAGAAAGAT 2286
 Db 1861 CTCACAGGTACGATGATCAAGATGAGAAATTTCCAGAGCTGACGAACTCAAGAAAGAT 1920

Qy 2287 TTGATTGGAATTGACAAATCTGTGTGTTCGCGGAGAGGAGTTCAATCCGTCTGGGAGCTC 2346
 Db 1921 TTGATTGGAATTGACAAATCTGTGTGTTCGCGGAGAGGAGTTCAATCCGTCTGGGAGCTC 1980
 Qy 2347 AGCAAGCTCTCGGGGAGAGGGGCTCCAGACGCGATGTCTTCTG----- 2391
 Db 1981 AGCAAGCTCTCGGGGAGAGGGGCTCCAGACGCGATGTCTTCTGATGTGAGAGAGC 2040
 Qy 2392 ----- 2391
 Db 2041 GCTTGTCTCAGAGAAATGTGTCACTGGGCAAGAAAGCTCCAGAGCCGAGCTCG 2100
 Qy 2392 -----TTCAAGACGTCGCTGATTAACAGAGCCGGGGGCTGACGCGCTCCAAATCAG 2442
 Db 2101 GCCACCAATTCAAGACGTCGCTGATTAACAGAGCCGGGGGCTGACGCGCTCCAAATCAG 2160
 Qy 2443 TTTAAAGTCCACGGGAGCTCCGCTCTATGAGATGACGATTTGAGAGAGAGCAAGAGAG 2502
 Db 2161 TTTAAAGTCCACGGGAGCTCCGCTCTATGAGATGACGATTTGAGAGAGAGCAAGAGAG 2220
 Qy 2503 TGGGGGGTGCCTCACTGCTGACCTTCGCGGGCCAGCGGAGTCAATCATCTGTGCGGCC 2562
 Db 2221 TGGGGGGTGCCTCACTGCTGACCTTCGCGGGCCAGCGGAGTCAATCATCTGTGCGGCC 2280
 Qy 2563 AGTTCTCGGTCGAGATGAGAGAGGTGTTGAGAGATCAAGATGGCCATTGACCTGGCG 2622
 Db 2281 AGTTCTCGGTCGAGATGAGAGAGGTGTTGAGAGATCAAGATGGCCATTGACCTGGCG 2340
 Qy 2623 GAGAAAGAGACAGCCCGCCCTGAGTTCTTGGCAGAGAGCCCGCTGACAAAGTCC 2682
 Db 2341 GAGAAAGAGACAGCCCGCCCTGAGTTCTTGGCAGAGAGCCCGCTGACAAAGTCC 2400
 Qy 2683 CTTGATGAAGCCACCGCGCTGACAGAGTCAAGAGATGAATGAGCCCTTGGCCACA 2742
 Db 2401 CTTGATGAAGCCACCGCGCTGACAGAGTCAAGAGATGAATGAGCCCTTGGCCACA 2460
 Qy 2743 TCGCTGAGAGCGGAGCCCGGACCGGCAACCAATGGTGAAGTGTGCTGGACCGC 2802
 Db 2461 TCGCTGAGAGCGGAGCCCGGACCGGCAACCAATGGTGAAGTGTGCTGGACCGC 2520
 Qy 2803 AACACAGGAGTCTCAATGATGACTTACAGCA-----TCGCAAGTGAAG 2844
 Db 2521 AACACAGGAGTCTCAATGATGACTTACAGCAATTCGAGTGAAGGCGCTTGTTCAG 2580
 Qy 2845 AATCAGTTGTCTGGAACCTGCTGAGGAAATTCAAAAACAGCAACGGGTGCGAAGCTG 2904
 Db 2581 AATCAGTTGTCTGGAACCTGCTGAGGAAATTCAAAAACAGCAACGGGTGCGAAGCTG 2640
 Qy 2905 TGGGTGTGTTCACAACTTGTGCGCTGTTCTTCAAAATTCACACAGGACATCATATCC 2964
 Db 2641 TGGGTGTGTTCACAACTTGTGCGCTGTTCTTCAAAATTCACACAGGACATCATATCC 2700
 Qy 2965 CTTGACAGCTGCTGTGCTCGGCTACTGCTCAACATCCCTGTAGTCCGAAATC 3024
 Db 2701 CTTGACAGCTGCTGTGCTCGGCTACTGCTCAACATCCCTGTAGTCCGAAATC 2760
 Qy 3025 CAGAAAGACTAGTGTTCAGCTGCACTTCAAGTCCAGCTTACTTCAAGGGGGAA 3084
 Db 2761 CAGAAAGACTAGTGTTCAGCTGCACTTCAAGTCCAGCTTACTTCAAGGGGGAA 2820
 Qy 3085 AGCGAGTACAGTTCGAAGGTGCA 3109
 Db 2821 AGCGAGTACAGTTCGAAGGTGCA 2845

RESULT 9
 BC004009
 LOCUS
 DEFINITION
 Mub musculus PERM, Rhogef (Arhgef) and plectstrin domain protein 1
 (Chondrocyte-derived), mRNA (cDNA clone IMAGE:3493093), partial
 cds.
 ACCESSION BC004009

VERSION	BC004009.1	GI:13278387
KEYWORDS	Mus musculus (house mouse)	
SOURCE	Mus musculus	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurungathi; Murioidea; Muridae; Murinae; Mus.	
REFERENCE	1 (bases 1 to 2632)	
AUTHORS	Klaunig, R.D., Collins, F.S., Wagner, L., Shamen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Butler, K.H., Schaefer, C.F., Hach, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marsina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ueda, T.B., Tohyuki, S., Caminci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.U., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wotley, K.C., Hale, S., Garcia, A.M., Gay, L.U., Hulik, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shcherchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schermer, A., Schein, J.E., Jones, S.J., and Marz, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
TITLE	12477932	
JOURNAL	2 (bases 1 to 2632)	
PUBMED	Strauberg, R.	
REFERENCE	Direct Submission	
AUTHORS	Submitted (28-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
JOURNAL	NIH-MGC Project URL: http://mgc.nci.nih.gov	
REMARK	Contact: MGC help desk	
COMMENT	Email: cgapbs-roman@nih.gov	
	Tissue Procurement: Gilbert Smith, Ph.D.	
	cDNA Library Preparation: Life Technologies, Inc.	
	cDNA Library Arrayed By: The I.M.A.G.E. Consortium (ILNL)	
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305	
	Web site: http://www.sbgc.stanford.edu	
	Contact: (Dickson, Mark) mcdgaxil@stanford.edu	
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.	
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov	
	Series: IRAC Plate: 8 Row: p Column: 11	
	This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.	
FEATURES	Location/Qualifiers	
SOURCE	1..2632	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/strain="Czech II"	
	/db_xref="taxon:10090"	
	/clone_image="IMAGE:3493093"	
	/tissue_type="mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMV."	
	/clone_id="NCL CGAP_Lu29"	
	/lab_host="DH10B"	
	/note="vector: pcMV-SPORT6"	
	<1..2632	
	/gene="Pmp1"	
	/note="synonym: Cdnp"	
	/db_xref="GeneID:223354"	
gene		

CDS		/db_xref="MG1:2446173"
	<1..1263	
	/gene="Farpl"	
	/codon_start=1	
	/product="Farpl protein"	
	/protein_id="AAH04009.1"	
	/db_xref="GI:13278386"	
	/db_xref="GeneID:223254"	
	/db_xref="MGI:246173"	
	/translation="VMLKNIQGMKLHLAHMKSEALEALETISIKGRRLHEPCRDPE LQKYCYPLNTFLRPLRLMYKHVLERLKHPPHNADFRDCRALAEITENVAOL HGNIKRENPOKHELKDLIGDNLVI PGSEFRLSLSLSGKLGQQRFENDQ LVTSRELTSNOFKYGGOLPIYGMTIEESEEWGPHCTLRCOROSIIVAASSRSLE MRKMEDIQMAIDLAKSNGPTPELLASSPPDNPSBDRTPAADSEDDLIASATSLSE ROAFRENTWVHCWKHSITSVMWDESIAVENQJSLDFRKNSNGOKLMVFTFN CLFEYKSHQSDHPLASLPILGYLSLTISEENIHKDVFVKLFKFSHYFFRAEEZYTH ERMMEVIRSATSSSASRAHIISHKESHLY"	
misc_feature	55..309	
	/gene="Farpl"	
	/note="RhogEP. Region: Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPases"	
	/db_xref="CDD:smart00325"	
	448..699	
misc_feature		
	/gene="Farpl"	
	/note="PPh. Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntrophins) and to be inserted within other domains. Mutations in Bruton's tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinaemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"	
	/db_xref="CDD:smart00233"	
	928..1218	
	/gene="Farpl"	
	/note="PH. Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntrophins) and to be inserted within other domains. Mutations in Bruton's tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinaemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"	
	/db_xref="CDD:smart00233"	
misc_feature		
	/gene="Farpl"	
	/note="PH. Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntrophins) and to be inserted within other domains. Mutations in Bruton's tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinaemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"	
	/db_xref="CDD:smart00233"	
ORIGIN		
Query Match	30.6%; Score 1054; DB 6; Length 2632;	
Best Local Similarity	89.0%; Pred. No. 3.5e-288;	
Matches 1138; Conservative	0; Mismatches 140; Indels 0; Gaps 0;	
Dy	1924 GTCAATGCAGGAACATTTCAGAGCATGAGAAGCACTGGGGCGTCACTGTGGAAGCAGC	1983
Dd	1 GTCAATGCTGAAAAACATTTCAGAGCATGAGAAGCACTGGGGCGTCACTGTGGAAGCAGC	60
Dy	1984 GAGGCCCTTGAGGCGCTTGAGAGATGAGATCAAGAAGCTCCGGGGCGCTGAGAACTTTGCG	2043
Dd	61 GAGGCCCTTGAGGCGCTTGAGAGATGAGATCAAGAAGCTCCGGGGCGCTGAGAACTTTGCG	120
Dy	2044 AGAAGACTTTGAGCTGCAAGAGGTGTGTTACTTAACCGCTCAACAACCTTCTCTGGGGCA	2104
Dd	121 CGAAGACTTTCAGCTGCAAGAGGTGTGTTACTTAACCGCTCAACAACCTTCTCTGGGGCA	180
Dy	2104 CTGACACGGGCTCATGACCTAACACAGGTCTCTGAGGGCGCTGTGCAACACACCACCGCG	2163
Dd	181 CTGACACGGGCTCATGACCTAACACAGGTCTCTGAGGGCGCTGTGCAACACACCACCGCACCA	240

QY 2164 AGCCAGCCGACCTTCAGGGACTGCCGACCCGCTTTGGCAGAGATCAGGAGATGCTGCA 2223
DB 241 AACCAAGCCCTCACTTCAGGGACTGAGAGCTGCGCTGGCGAGATCAAGAGATGCTGCGC 300
QY 2224 CAGCTCCAGCGTTCAGATGATCAAGATGAGAAATTTCCAGAAAGCTGCAGAACTCAAGAAA 2283
DB 301 CAGCTGACGGGACATGATTCMAAGTGAATCTTCAGAAAGCTGCAGAAAGCTGCAGAAA 360
QY 2284 GATTGATTTGGCATGCAATCTTTGGTTCCGGGAAAGGAGTTCACTCCGCTGGGAGC 2343
DB 361 GATCTGATCGGCAATTGCAATCTTTGATTCAGAGAAAGGAGTTCACTCCGCTGGGAGC 420
QY 2344 CTGAGCAAGCTTCGGGGAAGGGGCTCCAGACGGCATGTTCTTCCCTGTTCAAGAGCTC 2403
DB 421 CTGAGCAAGCTTCGGGGAAGGGGCTTCAGACGGCATGTTCTTCCCTGTTCAAGAGCTC 480
QY 2404 CTGCTATACAGAGCCGGGGCTGACGGCTCCAAATCAGTTTAAAGTCCAGGGGAGCTC 2463
DB 481 TTGCTGATACAGCCGGGACTGACAGCATCTAATCAAGTTTAAAGTCCAGGGAGCTC 540
QY 2464 CCGCTCTATGACATGACGATTTGAGAGAGCGAAAGCAGAGTGGGGGTGCCCACTGCTG 2523
DB 541 CCAGCTATGAGCATGACGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 2524 ACCCTCCGGGGCCAGCGGCGCATCTATCTGTGGCCGCGCATCTGCGTCCGAGATGAG 2583
DB 601 ACCCTTACGGGGCCAGCGGCGCATCTATCTGTGGCCGCGCATCTGCGTCCGAGATGAG 660
QY 2584 AAGTGGTTGAGAGATCCAGATGGCCATTCAGCTGCGGAGAGAGAGAGAGAGAGAGAGAG 2643
DB 661 AATGATGAGAGATCCAGATGGCCATTCAGCTGCGGAGAGAGAGAGAGAGAGAGAGAG 720
QY 2644 CTTGAGTTCTGCGCAGCAGCCCCCTGACAAAGTCCCTGATGAGCCAGCGGCT 2703
DB 721 CCGGACTCTCTTCCAGCAGCCCCCTGACAAAGTCCCGAGATGAGCCAGCGGCT 780
QY 2704 GACCAAGAGTCAAGAGATGACTGAGGCTCGCGGCATCTGCTGAGAGCGCCAGCGCCG 2763
DB 781 GACCAAGAGTCAAGAGATGACTGAGGCTCGCGGCATCTGCTGAGAGCGCCAGCGCCG 840
QY 2764 CACCGGGCCAAACAATGAGTGAAGTGTGCTGAGCAACCGCAACAGCGCTCTCATAGT 2823
DB 841 CACCGGGCCAAACAATGAGTGAAGTGTGCTGAGCAACCGCAACAGCGCTCTCATAGT 900
QY 2824 GACTTCAAGATGAGTGAAGTCAATGCTGAGAACTGCTGAGAAATTCAGAAATC 2883
DB 901 GACTTCAAGATGAGTGAAGTCAATGCTGAGAACTGCTGAGAAATTCAGAAATC 960
QY 2884 AGCAAGGGTGGAGAAAGCTGTGGTGTTCACAAATCTTCTGCTTCTTCAAA 2943
DB 961 AGCAAGGGTGGAGAAAGCTGTGGTGTTCACAAATCTTCTGCTTCTTCAAA 1020
QY 2944 TCACACACGAGCAATATCCCTTTGCGAGCTGCTGCTGCGGCTACTGCTCAGCATC 3003
DB 1021 TCACACACGAGCAATATCCCTTTGCGAGCTGCTGCTGCGGCTACTGCTCAGCATC 1080
QY 3004 CCGCTGAGTCCGAGAAATCCAGAAAGTACGTTTCAAGCTGACATTCAGAGTCCAC 3063
DB 1081 CCGCTGAGTCCGAGAAATCCAGAAAGTACGTTTCAAGCTGACATTCAGAGTCCAC 1140
QY 3064 GTCTACTACTTCAGGCGGAGAAAGCTGACAGTTCAGTTCAGAAAGTTCAGTTCAG 3123
DB 1141 GTCTACTACTTCAGGCGGAGAAAGCTGACAGTTCAGTTCAGAAAGTTCAGTTCAG 1200
QY 3124 AGTGACACAGCTCTGCTGCGAGACCCAGAGTTCAGCAAAAGTCTTGTGTAT 3183
DB 1201 AGTGACACAGCTCTGCTGCGAGACCCAGAGTTCAGCAAAAGTCTTGTGTAT 1260
QY 3184 TGATGCGCGGACCACTC 3201
DB 1261 TGATGCGCGGACCACTC 1278

RESULT 10
CQ725044 3997 bp DNA linear PAT 03-FEB-2004
LOCUS
DEFINITION
Sequence 10978 from Patent WO02068579.
CQ725044
ACCESSION
CQ725044
VERSION
CQ725044.1 GI:42285901
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kite, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 10978 06-SEP-2002;
PB Corporation (NY) (US)
FEATURES
source
Location/Qualifiers
1..3997
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 30.2%; Score 1039; DB 2; Length 3997;
Best Local Similarity 60.3%; Pred. No. 7, 2e-284;
Matches 1865; Conservative 0; Mismatches 1185; Indels 45; Gaps 7;
QY 69 GCCGACCCGAGATCCAGCTGGGGGCCCCGAAAAATTTGGGATAGTACTTGGAAAG 128
DB 150 GCAGACTGAGAGAGATGCGTTGGGTGCCAGACCCCTGGAGATTAGCACCCTTGAGCC 209
QY 129 TGGACAGAGCCGCCCCCAACCTTCAGAAAACTCGTCCATCAATCAATCCAGATGCT 188
DB 210 TGGGAGACTCTCTCTTCCAGAGATGCAAGAGAGACCTGCACTCAGATGAATAGCTGCT 269
QY 189 GGAATGACACCCAGAGAGCAATTTGAAGTTTCACAAAGCTCCTGGAGAGTGTCTGGA 248
DB 270 GGAATGACACCAATGAGAAATTTTGAATGAGCTTAAATGCAATGGCCAGATATTACTGAC 329
QY 249 TGCAGTTTGAACACACTGACCTGATGAGAGTGAATTTTGGCTGAGTTTCTCA 308
DB 330 ACAAGTGTGAAGCGTTTAACTGTGTGAATGTGCTCTTGGGATGAGATTTCAAA 389
QY 309 TCACAAAAGATCAGCGTGTGGTGGATCTCCAAAACCATTTGAAACAGATTAGAAG 368
DB 390 TACTCAGTCTCTAGATTTGGCTTGAACCTTGAACCACATCATTTAGCAAAATACAGAG 449
QY 369 GCCAAAGCACTGTTGTTTGAATTTGTGTGTAATTTCTTCGCGCTGACACACAACT 428
DB 450 GCCAAAGAAATGTGTGCTTGGCTTGGCTTAATAATTTTTCACCTGATCTGTGTAGCT 509
QY 429 CCAGAAAGAACTCAGAGTACTGTTGCGCGCTGCAAGTGAAGCAGAGATTGGCTCAAG 488
DB 510 ACAGAAAGAAATTAACAAGTACTTGTTCCTTGAACAATTAAGAGAGACTCTGTGAAGA 569
QY 489 CAGGTGAGGTGAATGACACAGCGCAGCTCTTGAATTTTCAACATGTGCAATCTGA 548
DB 570 GCGTTTGACTGTGCTGACACACAGCGGCTCTTCAAGTCCATCTCTGCAAGTGGGA 629
QY 549 GATTGGGAGTTTGTATGAAGCTTGGACAGAGAGCACTTAGCAAAAATTAATACATACC 608
DB 630 AATAGGAGATTACGATGAAACGCTGGAACGAGAGACCTCAAGTGAAGAGATATTGGCC 689
QY 609 TCAAGAGACGACATGAGAGCAAAATCGTGAATTTTCAACATACCATTTGAGCAAAAC 668
DB 690 TGGCCAGCAGCACTGCTTGAAGAAATTAATTAATTCATCAAGAGCAGTGGGCGAGC 749
QY 669 ACCAGCAGATGAGATTTCCAGCTCTTAGAGATTCGCCGCTGCGCTAGAGATATGAAAT 728

Db 750 AACTGCTGAGTCGGATTTTCAGAGTCTCGAATAATGCTCGAAGATTGGAATGTACGGCAAT 809
Qy 729 CCGGTTGCAACCGGCGCAAGAGACAGGGAAGGACGAAGATCAATCTGGCGGTTGGCAAC 788
Db 810 CAGATTTCACATGCTTCTGACAGGGAAGAACCAAGATTCAATGCGAGTTTCCACAT 869
Qy 789 GGGAAATTAGTGTTCAGAGGTTTCACTAAGATCAATGCTTCAACTGGGCGCAAGGTGCG 848
Db 870 GGGGTACTGATGTTCTCAAGGCGACACCAAAATCAACATTTCAACTGGTCCAGAGTCCG 929
Qy 849 GAAAGTACCTTCAGAGGAAGGCTTTCTCATCAAGCTCCGGCCGAGTCCCAATGTGCG 908
Db 930 TAAACTAACCTTCAGAGGAAGAAATTTCTTACAACTTCACAGAGGTTCTATGACC 989
Qy 909 GTACACAGATACCTTGGAAATTCCTGATGCGCAGTCCGGAATTTCTGCAAGTCTCTGAA 968
Db 990 TTAGCGGACACATTAAGAAATTTTGTGGGTAGTAGATGAATGTAAAGATCTTCTGGA 1049
Qy 969 AATCTGTGTGAACATCATGCTCTTCTTGAACCTTTTGAAGGCCCAACCAAGCCCA 1028
Db 1050 GATTGTGTGAGATACACACCTTTTGTAGACTTTTGAACCACTTACCAAAAGCAAA 1109
Qy 1029 GCCCGTCTCTTTAGCGGGGGGTTCATATTGGGTTCAGTGTGCGGACTCAGAGAGGT 1088
Db 1110 AGCGGTCTTCTTCAAGCGGGGCTCTCTTCAGATACAGTGAAGAACTCAGAAACACT 1169
Qy 1089 TCTCGACTATGTTTAAAGAGAGAGACATTAAGAAAGGTGACGTTTGAAGAGACACAGCA 1148
Db 1170 AGTAGATTATTTCAAGAGACAGTGAATGAAGAAATTCATATGAAGAAAGGACAGCA 1229
Qy 1149 GATTCTATCTATCCGAGGCTTGTTCACAGCTTACAGAACTGAATTCGAAAGTGTGA 1208
Db 1230 GACCCACAGTCC-----GTTTCAGCTGTGACGACAGCTACCA 1271
Qy 1209 GAGGTTCAGACAGACACAGCCTTACATTGGAAGAAAGTCCGAATCTTCAGAGGCGCA 1268
Db 1272 ACAGAGCATCTCATTTCCGAGGAAATTGAGACTCTGCTCCCTCCCACTTCAGCGAATGC 1331
Qy 1269 GAGCTCCCGGCGAGAAAGAAACGAAGTTTCGCGCGGAGAGCGGAGTCCACCCGAG 1328
Db 1332 CTTTTACTGTCTCTCTCCCTCCACTGTGATCCCTCTGGCTCGAGAGTTTAAAGACAG 1391
Qy 1329 CCCTGCGCGAGAGAAAGCCCGGGGTAAACAAGACGAGCGAGCGCTCGGCGCC 1388
Db 1392 CAGCAGCTCCCTACAGATCCCAAGGTTTCTTACGTCAAGAGTCCAGCTGACAGAGGCG 1451
Qy 1389 CACGGAAGAGAGAGAGAG---TGGTTAAGATAGAGACCCAGACAGTAAACCTTCAGCC 1445
Db 1452 CAGTGAAGAGTGGCTGAGAGGCCGGAACACCATCGGCCAGCCCTCGGGGCCCGCG 1511
Qy 1446 CCGGACGCAAGACAGAGCTCCCTGACTGGCAGTCTCTCACTTTCCGAGCTGTCTGTGA 1505
Db 1512 ACTCCAGCTGTGTCCAGGCTTTCCAGAAAGTCTTCAAGCTTCTCCCTCCAGCCGGA 1571
Qy 1506 CTGCGAGGGGGAGTGGCCCTGCAACGTGACTTGTCTCCCACTTGAAGCCCGGAC 1565
Db 1572 GAGCCCCCTGAGCTGAGCCCTGCAATTTGAGGTGCTT-----TGGGCCAGCTGA 1622
Qy 1566 CAGACAGGCTCTCTCTTATCAGCCCGGTGCTGAATGACAGGCTTCCCGCCCGAGCGA 1625
Db 1623 ACAGGGCTATCCCACTCTGAGCCCTGTCTCAGTGTGCTGCGGAGAGCGGGATGGA 1682
Qy 1626 CGATGAGATGAGAGGCGGAGAGAGATTTCCCACTGAATTAAGCCTTACATAGCTAA 1685
Db 1683 CTGCGAG---GAGCCAGACACAGGCGGTCTGAGAGAGAGGCTTACTTCAATAGTCA 1739
Qy 1686 GGAAGTGTATCAACGAGAGCAATATCTGAAGATCTCGAAATTATCACTTCTGTGTT 1745
Db 1740 AAGAGATTCTGCTACAGAAAGCAATACCTCAAGAGATTAGAAAGTTATACGTGTGTT 1799
Qy 1746 TCAAGACACAGTGAAGAGAGAGAGCCCATGCGGAGACAGCTGAAGAAAGTCTATATCC 1805
Db 1800 CCGACAGCAGTGTGTAAGAGAGAGCGCATGTGCTGACTGTGATGACCGTCTCTTCTC 1859

Qy 1806 GAATTTGAACTTTTGCACAAAATTCATATAATTTCTCAAGAAATGACAAAGCACT 1865
Db 1860 CAACATGATCCCATCTATGATTTCCACAGAGGCTTCTGGCGAGGTGAGACAGAGCT 1919
Qy 1866 TGCCTGTGGAGAGGCGCTCAATATGCCAATCAGAG---ATTACCAAGATGGCGA 1922
Db 1920 GGCACCTTGGAGAGGCGCTCCCAAGGCCACACAAAGGAGTCTACAGATGAGGGGA 1979
Qy 1923 TGTCACTGGAAGAACATTTCAAGGCGATGAGACACTGGCGGCTCACTGTGGAACACAG 1982
Db 1980 CATCTGCTCAGAGAACATGCGCAGTTAAAGAGTTTACAGCTACTTCCAAATACAGA 2039
Qy 1983 CGAGGCTTGGAGGCTCTGAGAAATGAAATCAAGAGCTCCGCGCGCTGAGAACTTGTG 2042
Db 2040 CGAGGCTCTTACAGAACTGGAAGAGCTAACCAAGCTGTAAAGATTGAGAGGAGTGA 2099
Qy 2043 CAGAGACTTTGAGCTGCAAGAGGTGTACTTACCTGCTCAACACACTTCTCTGCGGC 2102
Db 2100 CAGGAGTTTGAAGTCAAGAGGTCTGTACTTGTCTTCAACAGTTCTGTGTAAGGC 2159
Qy 2103 ACTGACCGGCTCATGCACTACAGCAGTCTGTGAGGCGGTGTCAAACACACCGGC 2162
Db 2160 CATCAGGAGCTGTGCACTACCGCTCTGTGCGCGCTGTGCGACATTAAGGCC 2219
Qy 2163 GAGCAGCGCCGACTTTCAGGGAATGCGGAGCGCTTTTGGCAGAGATCACAGAGATGAGC 2222
Db 2220 CGGCAACCATGACTACGCTGACTGCTGATGAGCGCTTGAAGCATCACAGAGGTACAC 2279
Qy 2223 ACAGTCCACGGTACAGATATCAAGATGGAAGATTTCCAGAACTGACGAATCAAGAA 2282
Db 2280 CACACTACAGACATTTCTCATCCGCTGGAACCTGACAGAACTTACAGAGCTGACAGC 2339
Qy 2283 AGATTGATGGAATGACAAATCTGTGTTCCGGAAGAGGAGTTTATCCGTCTGGGCG 2342
Db 2340 GGAAGTGTGGCATAGAGAACCTATGTCTCTGCGAGGAGTTTCACTGAGAGGCTG 2399
Qy 2343 CCTCAGAACTCTTGGGGAAGGGGCTTCAGACGCGCATGTTTCTGTTCAGAGAGT 2402
Db 2400 CTTTCAAGCTCACAGAAAGGGGCTGACAGAGAGAGATGTTTCTGTCTCAGATAT 2459
Qy 2403 CCTGTATATACAGAGCGGGGCTGACGGCTCCCAATAGTTTAAAGTTCACAGGCACT 2462
Db 2460 GTTGTGTACCAAGCAAGAGGATGACAGGACAGCACTTCGGATTCGGAGGCTCTCT 2519
Qy 2463 CCGGCTATGAGCATGATGAGTGAAGAGAGCAAGAGAGTGGGGGTGCCCATGCT 2522
Db 2520 TCCCTTCCAGAGCATGTGTGGAAGAAAGTGTAAACAGATGTGTCTTCCACATGTT 2579
Qy 2523 GACCTTCGGGGCCAGCGGAGTCCATCATGTGGCGGCAAGTTCTGAGTCCGAGATGGA 2582
Db 2580 CACCATATACGGGCTCAGAAACCAATGTTGTGCAACGACCTCGGCTGGAAGAAAGA 2639
Qy 2583 GAAGTGGTTTGAAGCATTCAGATGGCCATTGACTGTGCGAGAGAAACAGAGCCCGC 2642
Db 2640 GAAAGTGAATGTGACCTGAACTCCGCAATCCAAACACCAAGATGCGGTGAACAGGC 2699
Qy 2643 CCTGAGTCTTGGCAGCA-----GCCCGCTGACAAACAAGTCCCTGATGAAGCAC 2696
Db 2700 CCTGTCACTGCAAGGCGGACGTGTGCACTGTGTCCCGAGATCCCAAGAGATATC 2759
Qy 2697 CGGAGCTACAGAGTGAAGATGACTGAGCGCTTCGCGACATGCTTGAAGCGCA 2756
Db 2760 TCTG---GAGCAGAGGTACAGAAAGATGATGTGCGGGGTGTCCGAGCTCCCTGAGAGGACA 2816
Qy 2757 GGCCTCGACCGCGGCAACAAATGATGCAAGTGTGTGCTGCAACCGCAACACAGGTCTC 2816
Db 2817 TGGCAGGACCGGGCCAAACACCAATGACGTGTGCTGTAACCGGAACACAGCGTGTG 2876
Qy 2817 CATGTGACTTTCAGCATCGCAGTGAAGATCAAGTTGTCTGGAACCTGTGAGAAATT 2876
Db 2877 CAGGCAAGCACAGTGAAGTGTGTGAGAACAGCTTTCAAGATATCTGTGAAGAAATT 2936

QY	2877	AAAAACAGCAACGGGTGGCAAGACGTGGGTGGTGTCCAAACTTCGCTGTTCTT	2938
Db	2937	CAAAAACGTATGCTGGCAGAAAGCTCTGGGTGCTTTTACCACTTCTTTGTTCTT	2996
QY	2937	CTACAAATCACACCAGACAAATCATCATCCCTTGGCAGCTGCTCTGTCGCTACTGCT	2998
Db	2997	CTACAAAATCATCATGAGATGATCAACCACTGGCCAGCTCTCCGGCTGTGGGCTACAGGT	3056
QY	2997	CACCATCCCCCTCTAGTGTCCGAGAAATCTCCAGAAAGATCACTGTTCAGCTGCACCTTCA	3056
Db	3057	GAGATATCCCCAGGGAGGCCGATGGCATACAAAGACTATATTTTCAAGCTCCAGTTCAA	3116
QY	3057	GTCCACAGTCACTACTTCAGGGCGGAAGCGAGTACAGTTTCGAAAGGTGATGAGAGT	3116
Db	3117	ATCCACAGTCACTACTTCTTCGGGGGTGAAGACAGTACACATTTTGAAAGTGGATGAGGT	3176
QY	3117	GATCCGCAAGTCCACACAGCTCTGCTCGCGACCCC	3151
Db	3177	GATCCAGGGGCGCAGCAAGTCAAGCCGGAGGGGCC	3211

RESULT 11					
CO8895396	CO8895396	3997 bp	DNA	linear	PAT 05-NOV-2004
LOCUS					
DEFINITION	Sequence 16 from Parent EPI471153.				
ACCESSION	CO8895396				
VERSION	CO8895396.1	GI:55467577			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

REFERENCE	AUTHORS	TITLE	JOURNAL
1	Certa, U., Foser, S. and Meyer, K.	Transcriptional activity assay	Patent: EP 1471153-A 16 27-OCT-2004; F. HOFFMANN-LA ROCHE AG (CH)

FEATURES	Location/Qualifiers
source	1. .3997

ORIGIN

Query March	30.1%	Score 1037.4;	Length 3997;
Best Local Similarity	60.2%	Pred. No. 2.1e-283;	
Matches 1864; Conservative	0;	Mismatches 1186;	Indels 45; Gaps 7;

Qy	69	GCACACCCAGGATCAAGACTGGGGGGCCCCGGAAAAATTCGGGGATACGACTCTTGGAAAG	128
Db	150	GCACATCGACGGGATGCGCTGGGGTGCCCAAGCCCTGTGGGAAGTTACACCCCTTGAACC	209
Qy	129	TGACACAGAGCGCGCCCCAACACCTTCAGAAAACTCGTGTCCATCAAAATCCAGATGCT	188
Db	210	TGGCACAAGCTCTTGGCCCAAAATGCAAGAAACACCTGCACCTCAGAGTAAAGCTGCT	269
Qy	189	GGATGACACCCAGGAGGCGATTGAAAGTTCACAAAGAGCTCTGTGGAAAGGTGCGCTGGA	248
Db	270	GGACAAACACCATGGAATATTTTGACATTGAGCTTAAATGCGATGGCCAGATTAATCTAC	329
Qy	249	TGCAGTTGCACCAACCTCAACCTCGTGAAGGTACTAATTTTGACCTCGAGTTTCTGA	308
Db	330	ACAAAGTGGGAAGGTTTAAACCTGGTGAATGTGACTACTTCGGGATGGAGTTTCAAA	389
Qy	309	TCACAAAAGATCAAGGTGTGGCTGAGATCTCTTAAACCCATTGTGAAACAAGATTAGAG	368
Db	390	TATCAAGTCCATCGAGATTGTGGCTTGAACCATATGAACCATCATATAGCAAAATACAGAG	449
Qy	369	GCCAAAGCAGTTGTTGAAGTTGTGTGAAATCTTTTCGCGCTGACACACACAACACT	428
Db	450	GCCAAAGAAATGTGTGCTTCGCTCAGCTGATGAATATTTTTCACCTGATATCTGTGTACCT	509

Qy	422	CCAAAGAACTCAACAAGTACCTGTTCCGGCTGCAGGTAAAGAGAACTTGGCTCAGG	488
Db	510	ACAAGAAGAAATATACAAAGATATCTTGTTCCTTGCAACTTAAGAGAACCTGCTGANA	569
Qy	489	CAGGTTACCTGTATATACACACAGCGAGCTCTCTTGATTTTCCACATTGTGCAATCTGA	548
Db	570	GCGTTTACCTGTGCTBACACACAGGGGCCCTTCTACGTCCCATCTCCTGCAGTGGGA	629
Qy	549	GATTGGGGAATTTTGATGAGCCCTTGGACAGAGACACTTGAACAAAATAATATCATACC	608
Db	630	AATAGAGATTACGATGAACGCTGCACCGAGAGCACTCAAGTGAACGAGTATTGGCC	689
Qy	609	TCAGCAAGACGCACCTAAGACCAAAATGTGGAATTTTCAACCTAACCATTTGGACAAAC	668
Db	690	TGGCCACACACACTGCCTTGAAGAAATCTAAGAAATTCATCTAGAAAGACGTGGGCCAGAC	749
Qy	669	ACCAGCAGAAATCAGATTTCAGAGCTCCTAAGAGATTGCCGCTCGGTAGAGATGTATGAAT	728
Db	750	ACCTGCTGAAGTCGATTTTCCAGGTGCTCGAAATTTGCTCGAAAGTTGAAATGTACGGCAT	809
Qy	729	CCGGTTGCACCCGGCCAAAGACAGAGGAAGCAGAGATCAATCTGGCCGTTGCCAACAC	788
Db	810	CAGATTTCAATGSCCTTCTGACAGGAAAGAACCAAGATTCAACTGSCAGTTTCCACAT	869
Qy	789	GGGAATTCATGTGTTTTCAGGCTTTCACTAAGATCAATGCTTCAACTGGGCCAAGGTGCG	848
Db	870	GGGGTACTGTGTTCAGGGCACACCAAAATCTACATCTTCAACTGCTCAAGGTCCG	929
Qy	849	GAAGCTACGCTTCAAGAGGAAGGSCCTTCTCATCAACCTCCGGCAGATAGCAATAGTGC	908
Db	930	TAACTAAGCTTCAAGAGAAAGAAATTTCTTATCAAACTTCATCTCAAGAGTTCAATGAC	989
Qy	909	GTACCAAGATACCTTGGAAATTCCTGATGGCCAGTCGGGAATTTTGCAAAGTCTTCTGAA	968
Db	990	TTTACCAAGACACATTAAGATTTTGTGGGTAGTAAAGATGAATGTAAAGAACTTCTGAA	1049
Qy	969	AATCTGTGTGAACATCATGCTTCTTTAGACTTTTGAAGGCCCAACCAAGCCCA	1028
Db	1050	GATTTGTGTGAGATACACCTTTTTTAAGCTTTTGAACCACTTAAGCCAAAGGCAAA	1109
Qy	1029	GCCGCTCTCTTTAAGCCGGGGGTACATCTTGGTTGAGTGTGCGCATCAGAAAGCAGT	1088
Db	1110	AGCGCTTCTTTCAGCCGGGACTCTCTCTTCAGATACAGTGAAGAACTCAAGAAACACT	1169
Qy	1089	TCTCGACTATGTTAAAGAAAGAGACATTAAGAAAGTGCATTTTGAAGAAAGACACAGAA	1148
Db	1170	AGTAGATTATTTCAAGACAGTGAAGAAAGAAATTCATATGAAGAAAGCACACAGCA	1229
Qy	1149	GATTCAATTCATCCGAGACCTTGCTTTCACAGCCTTACAGAACTGAATCTGGAAGTCTGGA	1208
Db	1230	GACCCACACGTC-----GTTGAGCTCTGACTGACACCACTAACCA	1271
Qy	1209	GCAGTCTCAGACAGACACAGCCTTACATTTGGAAGAAAGTGTCCGAATCTCAGGGGGCCA	1268
Db	1272	ACAGAGATCTCATTTCCCGAGGAGATTGAGAACTCTGCTCCCACTTTCAGGAAATGC	1331
Qy	1269	GAGCTGCGCGCGAGAAAGAAACCGAAGTTTTCGCGCGGGAAGCCGGGATCGCACCCGAG	1328
Db	1332	CTTTTACTGCTCTCTCTCCCTCCTCACTCTGTGTCCTCTGTGCTGCAGATTTTAAAGACAG	1391
Qy	1329	CCCTGCGCCGAGGAAGAACCCCGCGGGTAAACAGACAGCGGAGCGAGCCGCTCTGCGCGC	1388
Db	1392	CAGAGCTCCCTCAGATCCCAAGGTTTCTTACGTCAAGATTCAGCTCAGAGAGGCG	1451
Qy	1389	CACGAGGAAGAGAGAGAG---TCTGTTAAGATAGAACCCACACAGATTAACCTCAGCC	1445
Db	1452	CAGTGAAGAGTGTGAGAGGCCCCACACACATCGGCGCAGACCCCTCGAGCCCCCGC	1511
Qy	1446	CCCGCAAGCCAAAGACAGAGCTCCCTGACTGCGCAGTCCCTCACTTTCGAGCTGTCTGAA	1505
Db	1512	ACTTCACAGCTGTGTCAGAGCCCTTTTCACAGAAAGTCTCTCAGCTTTCCTTCACAGCGGAA	1571
Qy	1506	CTGCGAGAGGAGATGGCCCCCTGCAACGTGACTTGTCTTCCAACTGTAGCCCCCGAC	1565

```

Db      1572 GAGCCCCCTGAGCTGAGCCCTGCAATTCAGGAGCCTT-----TGGCCCAAGCTGA 1622
Qy      1566 CAAAGAGGCTCTCCCTTATGAGCCCCGCTGATGATGACAGAGCTGCCCCCGACGGA 1625
Db      1623 AAGAGGCTATCCCACTCTGAGCCCTGCTCACTGATGCTGAGGAGCCGGAGTGA 1682
Qy      1626 CGATGAGATGAGGAGCGGAGAGAGATTCCTCACTGAATAAGCTGATAGCTTA 1685
Db      1683 CTGCGAG---GAGCCAGACAGAGGCGCTGAGAGAGAGGCTTACTTATGATCA 1739
Qy      1686 GGAAGTGTCTACACCGAGCGAAATATGTAAGATCTGAAATTAATCACTTCGTGTT 1745
Db      1740 AAGAGATTCGCTACAGAAAGAACTACCTCAAGAGATTAGAAGTATATACGTGTGTT 1799
Qy      1746 TCAGAGCAGAGTGAAGAGAGAGAGAGCGCAATGCCGAGAGAGCTGMAAAGCTATATCC 1805
Db      1800 CGCAGCGCAGAGTGTGAAGAGAGAGCGCAATGCTGCACTGATGACGCTGCTTCTTC 1859
Qy      1806 GAATTTTGAACCTTTGCACAAAATTTGATCTAATTTTCTCAAGAAATTTGAGCAAGCT 1865
Db      1860 CAACATCGATCCCATCTATGAGTTTCAAGAGGCTTCTGCGGAGGTGAGAGAGGCT 1919
Qy      1866 TGGCTGTGGAAAGGCGGCTCAATGCCAATCAAG---ATTACAAAGATGGGCGA 1922
Db      1920 GGCACCTGTGGAAAGGCGCTCCAAAGCCCAACAAAGGCGAGTCATCAAGATCGGGGA 1979
Qy      1923 TGTGATGCTGAAGAAATTCAGAGGCAATGAAGCACTGGGGGCTCACTGTGAAGCAG 1982
Db      1980 CATCTGCTCAGAAACATGCGCCAGTTAAAGAGTTTACAGATCTTCCAAAGACATGA 2039
Qy      1983 CGAGGCTTTGAGAGCCCTGAGAGATGAAATCAAGAGCTCCGCGGCTGAGAGACTTCTG 2042
Db      2040 CGAGGCTCTTAAGAACTGGAAGAGGCTACAAAGCTGTAAAGATTTGAGAGGAGTGA 2099
Qy      2043 CAGAGACTTTGAGCTGCAAGAGTGTGTTACTACCGCTCAACACTTCTCTGCGGCT 2102
Db      2100 CAGAGGTTTGAAGCTGCAAGAGGCTGTCTACTTGTCAACACTTCTCTGTGAAGCT 2159
Qy      2103 ACTGCAACGGGCTCATGCACTAACAAGAGGCTCCGAGGCGGTGTGAAACACACCGGCT 2162
Db      2160 CATCCAGCGAGCTGCTGCACTACCGCTGCTGCTGCGCCGCTATGAGACATTCAGCCC 2219
Qy      2163 GAGCCACGCGCACTTCAGAGGAGTCCGAGCGGCTTTGGAGAGATCAGAGATGAGTGC 2222
Db      2220 CGGGGACCATGACTAGCTGAGCTGACAGCGCCCTGAAGCATCAGAGAGTGAACAC 2279
Qy      2223 ACAGCTCCAGGATCGATGATCAAGTGAAGATTTCCGAAGCTGCAAGAACTCAAGAA 2282
Db      2280 CACACTACAGCACTTCTCATCCGCTGAGAACTGCAAGAGCTAACGAGCTGCAGCG 2339
Qy      2283 AGATTTGATTTGAGATGACATCTTGTGTTCCGGAGAGGAGTTGATCCGTCTGGGAG 2342
Db      2340 GAGCTGTGTGGGATGAGAACTCATTTGCTCTGAGAGAGGTTATCTCGTAGGGCTG 2399
Qy      2343 CCTCAGCAAGCTCTCGGAGAGAGGAGCTCAGAGCGCATGTTCTTCTGTTCAAGCACT 2402
Db      2400 CCTTCACAGGCTCACCAAGAGAGGCTGAGAGAGAGATGTTTTTCTGTCTCAGATAT 2459
Qy      2403 CCTGCTATACAGAGCGGGGGCTGACGCGCTTCATGAGTTTAAAGTCCAGGGAGCT 2462
Db      2460 GTTGTGTTACAGAGGAGAGGAGTTGACAGGACCACTTCGAGATCCGGAGCTTCT 2519
Qy      2463 CCGGCTCTATGAGCATGAGATTTGAGAGAGAGAGAGAGAGTGGGGGCTGCCCACTGCT 2522
Db      2520 TCCCTCTCAAGGAGTCTGTGTGAAGAGAGATTAACAGTGTCTGTCTCAGACTGTT 2579
Qy      2523 GACCCCTCCGGGCGCAGCGGCAATCATCTGTGCGGCAAGTTCTCGGTCCGAGATGA 2582
Db      2580 CACCATCTACGGGCTCAGAAAACATCTGTGTGAGCAGCAGCATCTCGGTGAGAAAGA 2639
Qy      2583 GAAGTGGGTTGAGGACATCAAGATGCCATTGACTGGCGGAGAGAGACAGACGCCCGC 2642

```

```

Db      2640 GAAGTGAATGTGACCTTGAATCTCGCATCCAGCAAGCAGAGTGGCGGTGACACGCG 2699
Qy      2643 CCTGAGTTCTTGGCAGCA-----GGCCCCGTGACAAAGATCCCTGATGAAGCCAC 2696
Db      2700 CCTTCATCTGCGAGGCGCAGCTGTGTGACTGTGCTCCCGCAGATCCCCCAAGAGTATC 2759
Qy      2697 CGCGGCTGACAGAGATGAGAGATGACCTGAGCGCTCGGCAATGCTGTGAGCGGCA 2756
Db      2760 TCTG---GAGCAGAGATCAGAAAGATGATGCTCGGGGTGTCCGAGAGCTCCTGGAAGGCA 2816
Qy      2757 GGGCCCGCAGCGGCGCAACATATGTGTGAGTGTGTGTGTCGACACCGCAACACAGCTCTC 2816
Db      2817 TGGCAGGACCGGGCCAAACACATGACAGTGTGTGTGTAACCGGAACACAGCGGTGC 2876
Qy      2817 CATGTGACCTTACAGCATCGCAGTGAAGATCAATGTGTGTGTAACCTGTGAGGAAAT 2876
Db      2877 CAGGCAAGCCAGAGTGAAGCTGTGAGAACAGCTTTCAAGATATCTGTGAAGAAATT 2936
Qy      2877 CAAAAACAGCAACGGGTGAGAGAGCTGTGGGTGTGTCACAACTTCTGCTGTGTT 2936
Db      2937 CAAAAACAGTATGGCTGAGAGAGCTGTGGGTGTGTCACAACTTCTGCTGTGTT 2996
Qy      2937 CTACAAATCAACACAGAGCAATCATCCCTTGTCCAGCGCTGCTGTGCTGCTATCCCT 2996
Db      2997 CTACAAACTATCATGAGATGACTACCACTGACGCTCCGCTGTGGGCTACAGGCT 3056
Qy      2997 CACATCCCTCTGAGTCCGAGAACATCAGAAAGACTACGTTGTTCAAGCTGCACTTCA 3056
Db      3057 GAGCATCCCGCAGAGAGGCGGATGAGTACACAAAGACATATGTTTCAAGCTCAGTTCA 3116
Qy      3057 GTCCAGCTCTACTACTTCAAGGCGGAAAGAGAGTACAGTTGCAAGAGTGAAGT 3116
Db      3117 ATCCACGCTTACTTCTTCCGGGCTGAGAGCAAGTACATTTGAAAGTGAAGT 3176
Qy      3117 GATCCGAGTGCACACAGCTCTGCTGCGACCCC 3151
Db      3177 GATCCAGGGGCGCAGACGCTCAGCCGGAGGAGCCC 3211

```

RESULT 12

```

AB018336          3997 bp      mRNA      linear      PRI 10-JAN-2004
LOCUS             AB018336
DEFINITION        Homo sapiens mRNA for KIAA0793 protein, partial cds.
ACCESSION         AB018336
VERSION           AB018336.1 GI:3882306
KEYWORDS
SOURCE
ORGANISM          Homo sapiens (human)

```

```

REFERENCE
AUTHORS           Ohara,O., Suyama,M., Nagase,T., Ishikawa,K. and Kikuno,R.
TITLE             Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
                   Prediction of the coding sequences of unidentified human genes. XI.
                   The complete sequences of 100 new cDNA clones from brain which code
                   for large proteins in vitro
JOURNAL           DNA Res. 5 (5), 277-286 (1998)
PUBMED            9872452

```

```

REFERENCE
AUTHORS           Ohara,O., Suyama,M., Nagase,T., Ishikawa,K. and Kikuno,R.
TITLE             Direct Submission
JOURNAL           Submitted (08-OCT-1998) Osamu Ohara, Kazusa DNA Research Institute,
                   Laboratory of DNA Technology, Yana 1532-3, Kisarazu, Chiba
                   292-0812, Japan (E-mail:cdna@fokkazusa.or.jp, Tel: +81-438-52-3913,
                   Fax: +81-438-52-3914)

```

FEATURES

```

source            1..3997
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /clone="hk05692"
                   /sex="male"

```



```

/issue type="brain"
/clone_lib="pBluescriptII SK plus"
/dev_sfile="adult"
1..3997
/gene="K1AA0793"
<115..3282
/gene="K1AA0793"
/notes="Start codon is not identified"
/codon_start=1
/product="K1AA0793 protein"
/protein_id="BA034513.2"
/db_xref="GI:40788367"
/translation="RMGEIEGTYSRLVLOTAGRLAQTQPVSVSLTEPQTLTPROMEKH
LHRLVLDNTMEIPDIEPKCDQVLLTQWKRLLVECOYPMENONTSYIMLEP
MKETIQRIRPKVULRIAPKPPRDPDGOQEBYTRYLPOLKROLLERLRLCATTT
MALTTHLLOSETGDETDIEDRLKNEYLYPGQHLKLEKLEFHQHQHQAETESDF
QVEIARKLEMYIIRPMASDRGTYKIQLASHMGVLPFGSTIKITPMNSKVRKSF
KRRFLIKLBEVHGYODTLEPLGSRDECKRMKICVEYHFPFLDLPKPKARAV
FPERGSFPRSGTKQQLVDYFMDSGMKRIPYRRHSKTHSVRLATLADPKOSIFP
EGLRTPASPSANAFYLSPSLTPSGLRKSTPSRPSRSLSTDPQVSVKPSAERBSAV
AGSDPSPASOPLRPPALOPRGLSTPSRPSRSLSTDPQVSVKPSAERBSAV
LSPVUSDGAGMDCBEPHKKVPADEAFIVKEILATERITLKLEVTITVPRSAV
VEDAMPATLMTLFSNIDPIYEFHGFLEVEQRLALWEGSPKATKSHQIIGIL
LRMRQLKEFTSYFORHDEVLELEKATKCKLEAVYEFELQKCYLPNTFLKP
IORLHYRLRLRCGYSPGHHDYADCDALKAITEVTTLQHLIRLENDOKLTBL
ORLVGIEBMLIARGREIRREGCKHLTKKGLQORMFPLSDMLLYTSKVGAGSHRI
RGILPIQGMIVESDNEKSPHCFITTAQKTIYVAASRLEKMKMLDLSHIOAK
SGSDIAPALPGRITVTRPPSPNEVSLQESBDADAGVNSLGHQOHRAHTIWHCW
YRNTSVRADHSAVENQSLGYYLRFKSHQWKQMVETNCLFPHYKHODYPLA
SLPLGSVSIPEADGINKDYVFKLPKSHVFPFAESKYTPERWMEVYOGASSSAG
RASIVQDGPQSSGLEGMVRKEE"

```

```

Query Match 30.1%; Score 1037.4; DB 5; Length 3997;
Best Local Similarity 60.2%; Pred. No. 2, 1e-283;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

```

ORIGIN

```

QY 69 GCCGACCCCGAGATCACTGGGGGCCCCGGAATAATGGGGATCAGTACTTGGAAAG 128
DB 150 GCAGATCGCAGGAGTGGCTTGGGTGCCAGACCCCTGGAGTTGACCCCTTGACC 209
QY 129 TGGCAGAAAGCCGCCCAACACCTTCAGGAAAATCTGTCCATCAAAATCCAGATGCT 188
DB 210 TGGGACAGACTCTTGGCCAGAAATGCAAGAAAGCAGCTGCACCTCAGATTAAGCTGCT 269
QY 189 GGAATGACCCAGAGAGCATTGAAGTTCACAAAGAGCTCTGGAGAGTGTCTGTGA 248
DB 270 GGAACAACCATGAAATATTGACATTGAGCCTTAATGCGATGGCCAGGTATTACTGAC 329
QY 249 TGCAGTTTGCACCACTCAACCTCGTGAAGGTAATTTTGGCTCGAGTTTCTGA 308
DB 330 ACAGAGTGGAAAGCTTTAAACCTGGAGATGAGTACTTGGGAGTGAAGTTTCAAA 389
QY 309 TCACAAAGAGTCAAGGTGTGGCTGATCTCTTAAACCCATTTGTGAAAAGATTAGAG 368
DB 390 TACTCACTCTTCAAGATTGGCTTGAACCTATGAAGAACCATCATTAAGGAATAGAG 449
QY 369 GCCAAAGCAGTGTGTTAAAGTTGTGTGAATTTCTTCCGCTGACCAACAACAAT 428
DB 450 GCCAAAGAAATGTGTGTGCTGCTAGCTGTAATAATTTTTCACATGATCTGTGAGCT 509
QY 429 CCAAGAAAGTCAAGGTAACCTGTTCCGCTGAGGTTGAAGCAGAGACTTGGCTCAAG 488
DB 510 ACAAGAAAGATTAACAGATATTGTTGCTTGGCAACTTAAGAGAGACCTGCTGAAGA 569
QY 489 CAGGTGACCTGTATGACACAGCGAGCTCTTGAATTTCAACATTTGCAATCTGA 548
DB 570 GCGTTTGACCTGTGCTGACACACAGCGGCTTTCTCAAGTCCCATCTCTGAGTCCA 629
QY 549 GATTTGGGATTTTGTGTAAGCTTGGACAGAGACCTTAAGCAAAAATTAATATACATACC 608
DB 630 AATAGAGATTACGATGAACGCTGACCGAGAGACCTCAAAAGTAAAGATATTGGCC 689

```

```

QY 609 TCAGCAAGACGACCTAGAGACAAATCGTGAATTTTCACTAACCAATTCAGCAAC 668
DB 690 TGGCCAGCAGCACTGCTTGAAGAGTACTAGAAATTCATCAGAAACAGTGGGCCAGAC 749
QY 669 ACCAGCAGATCAGATTTTCAAGCTCTCTAGAGATTTGCGGCTGAGATGTATGAAAT 728
DB 750 ACTGCTGAGTGGATTTTCAAGGTGCTCGAAATTTGCTCAAAAGTTGGAATGACGCGCAT 809
QY 729 CCGGTTGCAACCGGCCAAGACAGGGAAGGCAGAGATCAATCTGCGGCTGGCCAAAC 788
DB 810 CAAATTTCAATGCTTGTGACAGGGAAGAACCAAGATTTCACTGCGATTTCCACAT 869
QY 789 GGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCGCAAGTGGG 848
DB 870 GGGTGTACTGTGTTTCAAGGACACCAAAATCAACATTTCACTGATCCAAAGTCCG 929
QY 849 GAAGCTGACCTTCAAGAGAAAGCGCTTTTCATCAAGCTCCGGCCAGATGCCAATAGTC 908
DB 930 TAAACTAAGCTTCAAGAGAAAGATTTCTTATCAAACTTCAATCCAGAGTTTCATGAGCC 989
QY 909 GTTACCAAGATACCTTGAATTTCTGATGGCCAGTGGGATTTCTGCAATCTTCTGAA 968
DB 990 TTACCAAGACACATTAAGAAATTTTGTGGGTGTAGATGAATGTAAAGACTTCTGAA 1049
QY 969 AATCTGTGTAAACATCATGACCTTCTTGAATTTTGAAGAGCCCAACCAAGCCCA 1028
DB 1050 GATTTGTGTGAGTATCAACCTTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 1109
QY 1029 GCCCGTCTCTTTAGCGGGGGGTCACTATTTGCGTTCACTGTGCGATTCAGAGAGGT 1088
DB 1110 AGCGGCTCTTCAAGCGGGGGCTCTCTTCAAGATTCAGTGAAGAACTCAAGAAACA 1169
QY 1089 TCTGCACTATGTTAAAGAGAGACATTAAGAGTGCAGTTTGAAGAGACAGACGA 1148
DB 1170 AGTAGATTTATTTCAAGAGAGAGATGAAGAAATTCATATGAAGAAAGGACACGA 1229
QY 1149 GATTCATTTATTCGGAGCTTGTCTTCAAGCTTACAGAACTGAAATTCGAAAGTGTGA 1208
DB 1230 GACCCACAGTGC-----GTTCAGAGCTGTGATGTCAGACCTTACCAAA 1271
QY 1209 GCAGTCTCAGCAGAGACAGCCTTAATTTGAAGAGGTGCGAATTTCAAGGGGCGCA 1268
DB 1272 ACAGAGACTCTCAATTCGCCAGAGGATTTGAGAGCTCGTCCCTCCCATTTTCAAGCAATGC 1331
QY 1269 GAGCTCCCGGCGAGAGAAAGAACCAAGGTTTCCGCGGGAGCCGGGTTCGACCCGAG 1328
DB 1332 CTTTATCTGCTCTCTCCCTCAACTCTGATCCCTCTGCGCTGCGAGATTTAAAGACAG 1391
QY 1329 CCTTGGCGGAGAGAGAGCCCGGGGTAAACAGCAGGCGGACGAGCCCTCGGCGGC 1388
DB 1392 CAGCAGCTCTTCAAGATCTCCAGGTTTCTTACGTCAAGAGATCCAGCTGCAAGAGAGCG 1451
QY 1389 CACGAGAGAGAGAGAGAG--TGGTTAAGATAGAACCCAGCAGAGTAAACTTCAGCC 1445
DB 1452 CAGTGAAGCAGTGTGAGAGCCCGGACACACCATGAGGCCAGCCCTCGGGCCGCCGAC 1511
QY 1446 CCGCAGCCCAAGCAGAGCTCCCTGATCTGCAAGTCTTCACTTTCCAGAGTGTCTGAA 1505
DB 1512 ACTTCAGCTGTGTCAAGGCTTTTCAAGAGGTCTTCAAGCTTCTCTTCAAGCCGGA 1571
QY 1506 CTTCAGAGGAGAGAGTGGCCCTGCAACGTTGCTTCCCAACTGAGCCCGGACAC 1565
DB 1572 GAGCCCTGAGTGTGAGCTTGTCAATTAAGGTGCTT-----TGGGCCAGCTGA 1622
QY 1566 CAAAGAGGCTTCTCTTGTATGAGCCCGCTGTGAATGAACAGGCTTGCCTCCGAGCGA 1625
DB 1623 ACAGGAGCTATCCCACTCTGAGCCCTGCTCTCAGTATGTGTGGGAGAGCGGGATGGA 1682
QY 1626 CGATGAGATGAGGCGGAGAGAGATTTCCCACTGAATAAGCTTATTAATAGCTTA 1685
DB 1683 CTGCGAG--GAGCCCAAGCAAGAGGCTGTGAGAGCGAGGCTTATTAATAGCTTA 1739
QY 1686 GGAAGTGTACCAACGAGCAACATATCTGAAGATCTGAAATTAATCACTTCTGTGTT 1745

```


Db 1740 AGAGATTCGTGATACAGAGAAACATACCTCAAGATTTAGAAAGTTATATACCGTGTGGTT 1799
 Qy 1746 TCAGACACAGTGAAGAAAGAGAGCCCATGCGGAGCACTGAAAAGTCTCATATTTCC 1805
 Db 1800 CCGCAGCGAGTGTGTAAAGAGAGCCCATGCTGGACCTGTATACCTGTCTTTCTC 1859
 Qy 1806 GAATTTTGAACCTTTGCACAAAATTTATCTAATTTTCTCAAGGAAATTTGAGCAACGACT 1865
 Db 1860 CAACATCGATCCCATCTATGATGATTTCCACAGAGCTTCTGCGGAGAGTGAAGAGAGCT 1919
 Qy 1866 TGGCCTGTGGAGAGCGCGCTCAAAATGCCAATCAGAG--ATTACAAAGAAATCGGCGA 1922
 Db 1920 GGCACCTGTGGAGAGGCGCTCCAAAGCCACACAAAGAGAGTCAACAGAAATCGGCGA 1979
 Qy 1923 TGTATGCTGAGAGAAATTCAGAGGCAATGAAGCACTGGGCGCTCACCCTGTGAAGACAG 1982
 Db 1980 CATCTGCTCAGGAAACATGGCCAGTTAAAGAGTTTACCAAGCTACTTCCAAAGACATGA 2039
 Qy 1983 CGAGGCTTGAAGCGCTGAGAAATGAAATCAAGAGCTCCGCGGCTGAGAACTTCTG 2042
 Db 2040 CGAGGCTTGAAGCGCTGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2099
 Qy 2043 CAGAGACTTGAAGCTGAGAGAGTGTATACCTACCGCTCAACACCTTCTCTGCGGCGC 2102
 Db 2100 CAAGAGTTTGAAGCTGAGAGAGTGTATACCTACCGCTCAACACCTTCTCTGCGGCGC 2159
 Qy 2103 ACTGACCGGCTGATGATCACTAAGAGAGTCTGAGAGCGCTGTGAAACACACCGCGC 2162
 Db 2160 CATCCAGCGGCTGCTCCTACCTACCGCTGCTGCGCGCTGATGAGAGCTTACAGCGCC 2219
 Qy 2163 GAGCCAGCGGCTGCTCCTACCTACCGCTGCTGCGCGCTGATGAGAGCTTACAGCGCC 2222
 Db 2220 CGGCGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2279
 Qy 2223 ACAGCTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2282
 Db 2280 CACACTACAGACATCTCATCGGCTGAGAGAGCTGAGAGAGCTAACGAGCTGAGAGC 2339
 Qy 2283 AATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2342
 Db 2340 GAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2399
 Qy 2343 CCTCAGCAAGCTCTCGGAGAGAGGAGCTCAGAGAGGCTATGTTCTCTGTTCAACAGAGT 2402
 Db 2400 CCTTCAAGGCTCAGAGAGAGGAGCTCAGAGAGGAGTGTGTTCTCTGTTCTCAGATAT 2459
 Qy 2403 CCTGCTATACAGAGCGGAGGCTGAGCGGCTTCAATCAGTTTAAAGTTCAAGGAGCT 2462
 Db 2460 GTTGTGATACAGAGAGGAGTTGAGAGAGCAAGCTTCCGAGATCCGAGGCTCTCT 2519
 Qy 2463 CCGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2522
 Db 2520 TCCCTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2579
 Qy 2523 GACCTCTCGGAGAGAGGAGCTCAGAGAGGCTATGTTCTCTGTTCAACAGAGT 2582
 Db 2580 CACCATCTAGCGGCTCAGAGAGAGGAGCTCAGAGAGGAGTGTGTTCTCTGTTCTCAG 2639
 Qy 2583 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2642
 Db 2640 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2699
 Qy 2643 CCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2696
 Db 2700 CCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2759
 Qy 2697 CCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2756
 Db 2760 TCTG---GAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2816
 Qy 2757 GGGCCGCAACCGGAGCAACAAATGATGATGATGATGATGATGATGATGATGATGATG 2816

Db 2817 TGGCCAGACCGGAGCCAAACACCAATGACAGCTGTGTGTGATCCGGAACACCGAGCTGTC 2876
 Qy 2817 CATGTGATCTTACCATTCGACATGAGATGATGATGATGATGATGATGATGATGATGATG 2876
 Db 2877 CAGGCGAGCCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2936
 Qy 2877 CAAAACAGCAACCGGAGGAGAGAGCTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2936
 Db 2937 CAAAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2996
 Qy 2937 CTACAAATCACACAGACATATATCTCCCTGACAGCTGCTGCTGCTGCTGCTGCTGCT 2996
 Db 2997 CTACAAATCTCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 3056
 Qy 2997 CACATCTCCCTGT 3056
 Db 3057 GAGCATCTCCAGGAGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATG 3116
 Qy 3057 GTCCACAGCTCTACTACTTCAAGGCGGAAAGCAGATGACAGTTGAAAGGTGATGATG 3116
 Db 3117 ATCCACAGCTCTACTCTTCCGCGCTGAGAGCAAGTATGATGATGATGATGATGATG 3176
 Qy 3117 GATCCGACAGTGCACACAGCTGCTGCTGCGACCCC 3151
 Db 3177 GATCCAGGAGGCGCAGACAGCTGACCGGAGGAGCCC 3211

RESULT 13
 BC009153
 LOCUS
 DEFINITION
 Mus musculus PERM, RhoGEF and pleckstrin domain protein 2, mRNA
 (cdna clone MGC:6304 IMAGE:2655209), complete cds.
 BC009153
 VERSION
 KEYWORDS
 MGC.
 Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Mus.
 1 (bases 1 to 3719)

REFERENCE
 AUTHORS
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Uedl, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywicki, M.I., Skalka, U., Smillius, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 2 (bases 1 to 3719)
 Strausberg, R.
 Direct Submission
 Submitted (05-JUN-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov

REMARK
 COMMENT

Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 5 Row: m Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA g1: 21704021.

Location/Qualifiers

1. 3719

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N-3"

/db_xref="taxon:10090"

/clone="MGC:6304 IMAGE:2655209"

/issue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month

old mouse. Taken by biopsy."

/clone_lib="NCI CGAP_Mam2"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

1. 3719

/gene="Farp2"

/note="synonym: Flr"

/db_xref="GeneID:227377"

/db_xref="MGI:2385126"

99. 3296

/gene="Farp2"

/codon_start=1

/product="FERM, Rhogef and plectretin domain protein 2"

/protein_id="AA09153.1"

/db_xref="GI:14318719"

/db_xref="GeneID:227377"

/db_xref="MGI:2385126"

/translation="MGEIEGTVALPTSGRLGGOTAIGVTLPEPQSLSPRQEKM
 RIRVKLIDSTVELFDIEPKDGOVLNWKHNLTECYFGLEPQSNYSWIKEM
 KPIROVRKPKAVLRLAVKFPDPDQLOEYTRILPALQKRLIEBLTTNTA
 ALIHLQSEIGDYDETLDRHLKANEYLPNQESLEKILDRHQHTQTPAESDFQ
 VLEIAKLMEYGRFPMASDREGTKINLAVSHGVLPQGTTFINFNMSKPKSPK
 RKRFLLKHPVEVGPYODLEFLGSRDECKNFKICEYHTFRLSDPKPKAKPV
 FSRGSSFRYSGRKQOLVDYKDGKMKRIPEYRHSKTRSLHALTVDPKQSVTD
 GLRTASISSANYSFPPRSSLSPPGLPMLKDSLSLVDPQAPVKSTAAABSGPS
 SSGSPSTQSAHLPRPVLRPGPFSMDSPOPSLSKSHSLDCEIQLALSTAGAS
 PVLSPVLSGAGTARMQNEQKHKHMEDEAVFAIEILATEKTYLQDEIVTWERS
 ILLRNKRLKEPTSYFQRDHDEVLLEKATKCKLEAVKEBELQKVCCLPLNTFL
 KPVORLVHRLILSLCAHSPGRADYDEALKAITEVTEILOSLRLEMLQKLT
 ELORDLVGENTLARGREPRREGCKLTKKGLQRMFLPSMLYTSKMSAGAHF
 RIRGFLPLMGMLVESENEKSVLACFTTAAQKTIIVASTRIEKKMKQDNLAAIOA
 AKTIGSPVLLGGPVYTRTPRSSDEVSLSESDGKNGSLGNSQHNATHMVC
 YNTGSRADSHAVENOLSGYLLRKFGKNGWGLVVFTECLFFYTHODYDLA
 SLPLGYSVSLPREADSIHNDYVLFKQFSHYFFPAESKYTERFWMVDVYKRASSSPG
 RPSFTQDSHHSPLGEAIREKEACPSPLDKNL"

ORIGIN

Query Match 29.6%; Score 1019.2; DB 6; Length 3719;
 Best Local Similarity 59.1%; Pred. No. 3.2e-278;
 Matches 1867; Conservative 0; Mismatches 1263; Indels 30; Gaps 6;

QY 70 CCGAAGCCGAGTACGACTGGGGGCCCCGGAATAATTCGGGATCGACTGAGAACG 129
 DB 132 CCAAGTTGAGGACACGCTGGGTGGCCAGACCGCATTTGGAGTGTACCTCGAGCCA 191
 QY 130 GCAGAGAGCCGCCCAACACCTTCAGAAAACCTGTCATCAAAATCCAGATGCTG 189

DB 192 GAGCAGAGCGCTGTACCCAGATGACAGAGAACACATCGCATCAGAGTGAAGCTGCTG 251
 QY 190 GATGACACCCAGAGGACATTTTGAAGTCCAAAGAGCTCTGGGAAGGTGCTGTGAT 249
 DB 252 GACAGACACTGTAAAGTATTTTGCATCGAGCCGAAGTGAACGGGAGGTCTTACTGACT 311
 QY 250 GCAGTTTGACACCACTCAACCTCGTGAAGGTGATATTTTGGCTCGAGTTTCTGTGAT 309
 DB 312 CAGGTGTGAAGCATTTGATATCTCATTTGAATGTGACTACTTTGTGTGAAGTTCAAGAA 371
 QY 310 CACMAAAGATCAAGGTGTGCTGATCTCTTAAACCAATGTGAAACAGATTTAGAACG 369
 DB 372 GTCCAGTCTTACATGATTTGGCTTGAACCTATGAACCATCTTAGGACAGTACGAAG 431
 QY 370 CCAAGCAGCTTGTATTGAATTTGTGTGAATTTCTTCCGCTTACACACACAACTC 429
 DB 432 CCAAAAAAGCGGTGCTTGCCCTGGAGTAAATTTTTCCCGCTTACCTGTGCTGACTG 491
 QY 430 CAAGAAGACTCAAGTATACCTGTTGCGCTGACAGTGAAGCAGACTTGGCTCAAGGC 489
 DB 492 CAGAGAGTACACAGAGTACCTGTTGCTTGCCTTGCACCTCAAGAGACCTCTCGAGGAA 551
 QY 490 AGGTTGACGTGTATGACACAGCGAGCTCTTGTATTTACATGTGTGCAATGTGAG 549
 DB 552 CGCCTGACCTGACCGGCAACTGCAAGCCCTTCTCATATCCACTTGTGAGTGGAA 611
 QY 550 ATTGGGGAATTTTGTATGAAGCTTTGACAGAGACATTAGCAAAAAATTAATATCACT 609
 DB 612 ATCGGAGATTAATGATTAACCCCTGATTCAGAAACACCTCAAAAGCAATAAATCTTGC 671
 QY 610 CAGCAAGACGACTAGAGCAAAATCGTGAATTTACCATTAACCATTTGAGCAAAACA 669
 DB 672 AACCAAGAGAAATCTTAAAGAAAGTACTAGACTTCAATCAGAGGACACGGGCGAGCT 731
 QY 670 CCAGCAGAAATCAATTTCCAGCTCTAGATGATTCGCGCTGAGAGTGTATGGAATC 729
 DB 732 CTGACAGAGTCAATTTCCAGGCTTGTGATGACAGAAAGCTGAAATGTATGCAATC 791
 QY 730 CGGTTGACCCCGGCAAGACAGGAAAGGACAGAAATCAATTTGCGCTTGCACACAG 789
 DB 792 AGGTTTCAATGCTTCTACACAGAAAGGACCAAAATTAATCTACAAATTTCTCACATG 851
 QY 790 GGAATCTAGTGTTCAGGATTTCACTAATGATCAATGACCTTCAACTGGGCAAGGTGCG 849
 DB 852 GGTGTCTGTGTTCAGGATTTACACCAATTAATCACTTCACTGCTGCTCAAGTCCGT 911
 QY 850 AAGCTAGCTTCAAGAGAGCGCTTCTCATCAAGCTCGGCAAGTCCCAATAGTGGC 909
 DB 912 AACCTAAGCTTCAAGAGAAAGGTTTCTTATCAAACTCACCTGAGGTCCATGGGCC 971
 QY 910 TACCAAGATACCTTGAATTTCTGATGCGCAGTGGGATTTCTGCAAGTCTTTTGAAA 969
 DB 972 TACCAAGACACGTTAAGATCTTGTGGGTAGCAGAGTAAATGTAAAGACTTGTGAAG 1031
 QY 970 ACTGTGTTGAACATGATGCTCTTTAGACTTTTGAAGAGCCAAAGCCAAAGCCCAAG 1029
 DB 1032 ATATGTGTGAGTACATCACTTTTAACTCTGACACAGCTTAAAGCCAAAGCAAAG 1091
 QY 1030 CCCGTCTCTTTAGCCGCGGATCATCATTTGCTGATGATGATGCTGAGACTGAGAGCAG 1089
 DB 1092 GGTGTCTTCTTACGCGGAGGTCTCTCTCAATACAGTAAAGTAAAGTCAAGAAACAATA 1151
 QY 1090 CTGACTATGTTAAAGAGAGACATTAAGAAAGTGCAGATTGAAAGAGACACAGCAAG 1149
 DB 1152 GATGATTAATGAAACAGTGTGAATGAAGAAATTTCCATAGAAAGAGGACACATTAAG 1211
 QY 1150 ATTC--ATTCTATCCGAGACCTTGCTTACAGCTTACAGACTGAGACTGAATTCGAAAGTCTG 1206
 DB 1212 ACTCGACATCTCTTATGCTGTGACTGTAGATCTGCTTAAACAGAGCTCTCTTACAC 1271
 QY 1207 GAGCAGTCTCAGACAGACACAGCCTTACATTTGAGAAAGGTGCGGAATCTCAGAGGGGC 1266
 DB 1272 GATGCTGTAGAACTTCTGCTCTGCTCTTCAAGCAAAATGTCTCTTATATCC----- 1324

QY 1267 CAGAGCTGCGGCGAGAGAAAGAACCGAAGGTTTCCGCGGGAGCGGGGTCGACCCG 1326
 Db 1325 ---ACCCCTAAGTTCTTTTGTCCCTCTGCTGTCGCCAATTGAGAGACAGACGACG 1381
 QY 1327 AGCCCTGCGCGGAGAGAAAGCCCGCGGGTAAACAGACGCGGAGCGAGCCGCTCGCG 1386
 Db 1382 CTCCTTGTGTGATCCGCGAGGCTCCCGTCAATCAAGACACAGCAGCAGAGAGGACGG 1441
 QY 1387 CCCACGAGAGAGAGAGAGGTCTTAAAGATAGACCCAGCAGAGTAAACCTACGCC 1446
 Db 1442 ACCA--TCATCATCTTACCGCCCGCCAGACACAGTCGCCCCATCTCCCTGACCCCGGTG 1499
 QY 1447 CCGCAGCCAAAGACAGAGTCCCTGACTGGGAGTCTCACTTTCGAGCTGTCTGTGAC 1506
 Db 1500 CTCGGCTGTGTCCAGGCTTTTCTATGATATGCTCTCAGCTTCTCCCTCCAGCCTGAAG 1559
 QY 1507 TCGCAGGGGGAGATGGCCCTGCGCAAGTACCTTGTCTCCAACTGAGCCCGACACC 1566
 Db 1560 AGCAGCTGAGCCTGTGCTGCTGAGCTTCAAGCCGCG-----CTGAGTACAGCTGAG 1610
 QY 1567 AAGCAGGCTCTTCCCTTGAATCAGCCCGCTGTGAATGACACGACCTGCCCCGAGAGAC 1626
 Db 1611 CAGGGTGATCCCTGTGTCTAGCCCTGTGTCTAGGGCGCTGTGACAGCCAGATGAGC 1670
 QY 1627 GATGAGATGAGGGCCGAGAGAAAGATTCCCACTGATAAAGCGTAACTTCACTAGCTAG 1686
 Db 1671 AACCAAGAGAGAGAGAAACACAGACATGCGCAAGAGATGAGGCCATTTTCAATAGCAAG 1730
 QY 1687 GAAGTGTCTACACCGAGCGAAATATCTGAAGATCTGAAAGTTATCACTTGTGTTT 1746
 Db 1731 GAGATTCTGCTACAGAACGAACTTATGAGAGATTTAGAAAGTTATCACTGTGTGTTT 1790
 QY 1747 CAGAGCAGGTGAGCAAAAGAGACCGCATGCGGAAAGCACTGAAAAGTCTCATTTCCG 1806
 Db 1791 CGAGAGGTCTGATCAAGAGAGAGGCGCATGCGCAGCCCTGATGGCCCTTCTTCTCC 1850
 QY 1807 AATTTTGAACCTTTGACAAATTTCACTAATTTCTCAAGAAATTTAGCAACGACTT 1866
 Db 1851 AACATTGATTCAGTCTACAGATTCACAGAGCTTCTTCAAGAGTGAACAGAGGCTG 1910
 QY 1867 GCGCTGTGGAGAGCGCGCTCAATGCGCA--AATCAGAGATTTACCAAGAAATCGGCGAT 1923
 Db 1911 GCACTCTGGAGAGGCGCTCCAGTGCCTTAATAAGGTGATCAACAGGAAATCGGGGAC 1970
 QY 1924 GTCATGTCTGAAGAACTTTAGGGCAAGACCTTGGCGGCTCACTGTGTGAAGCAGAC 1983
 Db 1971 ATCTCTCTCAGAGAACTGCGTCACTTAAAGAAATTTTACCTTCAAAAGACAGAT 2030
 QY 1984 GAGGCTTTGGAGGCGCTGGAGAAATGGAATCAAGAGCTCCCGGCGGCTGGAGAACTTCTGC 2043
 Db 2031 GAGGTCTTAACAGAACTGGAAGAGGCAAAACATGTAAGAAAGTGAAGGCAAGTCTAC 2090
 QY 2044 AAGACTTTTGAAGTGAAGAGTGTGTTACTACGCTCAACACTTCTCTCTCGCGCCA 2103
 Db 2091 AAAGAGTTTGAAGTCAAAAGGTCTGTCTGTCTCAACACTTCTCTCTGAAAGCCC 2150
 QY 2104 CTGCAACGGGCTATGATCAAGCAGGTCTGTGAACGGGCTGTGCAAAACACACCGCGG 2163
 Db 2151 GTCCAGAGGCTGATCACTACGCTGTGTGTGAGCGGCTGTGTCTCACTCTCTCT 2210
 QY 2164 AGCCACGCGGACTTCAAGGACTGCGAGCGCTTTGGCAGAGATCAAGAGATGTGTGCA 2223
 Db 2211 GGGCACCAGGACATATGCCAGTCCATGTAGGCACTGAAGCCATCAAGAAAGTGAACCAAC 2270
 QY 2224 CAGCTTCACGCTGATCGATGATCAAGATGAGAAATTTCCAGAGCTGACCAACTCAAGAAA 2283
 Db 2271 GAGCTTCACCAAAAGCTTACCGCGCTGAGAAAACCTTCAAGAAATTTGACGAGCTACAGCA 2330
 QY 2284 GATTGATTTGGCATTTGACATCTGTGTGTTCCGGGAAAGGAGATTCATCCGTCTGGGCAAG 2343
 Db 2331 GACCTGTGTGTGTAGAAAACCTCATTTGCTCTGGGAGGAGATTTATCCGTGAGGGCTGC 2390

QY 2344 CTCAGCAAGCTCTCGGGAGAGGGGCTCCAGCAGCGCATGTTCTTCTGTTCAACGACGTC 2403
 Db 2391 CTGCAAGAGCTCATCAAGAGGGCTGTGACAGAGAGATTTTTTCTGTTCTCAATATG 2450
 QY 2404 CTGCTATACAGAGCCGGGGCTGACGGCTTCATCACTTTAAAGTCCAGGGCAGCTC 2463
 Db 2451 TTACTGTATACAGCAAAAGTGTCAAGAGGCAAGTCACTTCCGAGTCCGTGGCTTCTT 2510
 QY 2464 CCGCTCTATGGCATGAGATTGAGAGAGCGAAGCAGTGGGGGGTCCCCACTGCGCTG 2523
 Db 2511 CCATCCGTGTGATGTGTGTAGAGAAATGATGATGATGATGATGATGATGATGATGATG 2570
 QY 2524 ACCCTCCGGGGCAGCGGCAAGTCCATCATCTGTGCGGCGCAAGTTCCTGTGCGAGATGAG 2583
 Db 2571 ACCATCTATGACGCTCAGAAAACATTTGTGTAGACGACGACACTCGGCTAGAAAAGGA 2630
 QY 2584 AAGTGGTTGAGAGATCAAGATGCGCATTTGACTGTGGGAGAGAGACAGACGCCGCC 2643
 Db 2631 AAGTGAATGACAGAACTTAAATGACCAATCCAGACGCAAGACTATGATGATCTACACC 2690
 QY 2644 CCGTGAATCTGGGCGAGCGCCCTGACAACAAGTCCCTGATGAAAGCCAGCGGCT 2703
 Db 2691 CCAATGCTGTGG--GAGGCCGGGTATTAATCTGTACCCCTTGAATCTTGTGATGAATG 2747
 QY 2704 GACCAAGATGAGAGATGACCTGAGCGCTCGCGCATCATGCTGTGAGCGCCAGCCCG 2763
 Db 2748 TCTGTGAAGATCAAGAAATGATGTGAGAGAAACCGGGGCTCCCTGAGGGGAAACAGCCAG 2807
 QY 2764 CACCGCGGCAACATATGTGTGATGCTGTGTGCGCACCGCAACAGAGGTCTTCCATGTG 2823
 Db 2808 CACCGGGCAATACAAATCATCATGTGTGTGTGATCCGTAAATCAAGGTGTGTCAGACCA 2867
 QY 2824 GACTTACATTCGAGTGAAGATCAATGATCTGTGGAACCTGTGAGAAATTTCAAAAAC 2883
 Db 2868 GACCAAGTGCAGCTGTGTGAGAACAGCTTTCAAGATTTCTGTGAGAAATTTCAAGAAC 2927
 QY 2928 AGTAATGTGTGGCAGAACTCTGGGTGTCTTTAACCACTTCTGTGTCTTATATAA 2987
 Db 2944 TCACACAGACATATCCCTTGTGACGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 3003
 QY 2988 AACATTCAGAGATGATTAACCCCTGTGCGCAGCTCCCACTATGGTTACAGGTGAGCTC 3047
 Db 3004 CCGCTGAGTCCGAGAACATCCAGAAAGACTACGTTCAAGCTGCACTTCAAGTCCAC 3063
 QY 3048 CCCAGGGAGGCTGACAGATCAAGAAATATGTTCAAGCTCAATTTCAATCTCAC 3107
 Db 3124 AGTCCACCAAGTCTGTGCTCGGACCCCAAGTGTGAGCCCAAGAGATCTTGTGTAT 3183
 QY 3168 AAGGCGACAGCTCAACGAGAGACCCCAAGTTTCACTCAGAGCTGTGCTCAATCATCT 3227
 Db 3184 TGATGCGCGACACACTGTTCCGAGTGTGCTTCC 3223
 QY 3228 CAGGGCTGTGAGGAGAGATCAAGAAAGAGGAGTACC 3267

RESULT 14
 BC027077 2431 bp mRNA linear ROD 21-OCT-2003
 LOCUS BC027077
 DEFINITION Mus musculus PERM, RhogEP (Artyge) and pleckstrin domain protein 1 (chondrocyte-derived), mRNA (CDNA clone IMAGE:5010682), partial cds.
 ACCESSION BC027077
 VERSION BC027077.1 GI:20071584
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE AUTHORS Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Mus.

1 (bases 1 to 2431)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klauer, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, A.I., Wang, J., Heide, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schemetz, J.E., Brownstein, M.J., Umedin, T.B., Toshlyuk, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, J., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherzer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

TITLE Strausberg, R.
JOURNAL Direct Submission
PUBMED Submitted (04-APR-2002) National Institutes of Health, Mammalian
REFERENCE Gene Collection (MGC), Cancer Genomics Office, National Cancer
AUTHORS Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
JOURNAL USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsesed, H.,
Kowls, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nannavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAK Plate: 45 Row: c Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency OR analysis.

FEATURES location/Qualifiers
source 1. 2431

gene /organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/feature="IMAGE:5010682"
/feature="Mammalian tumor, MMTV-LTR/INT3 model, 5 month old mouse, Taken by biopsy."
/clone_1lb="NCI CGAP_Mam2"
/note="Vector: PCMV-SPORT6"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
/gene="Farpl1"
/codon_start=1

/product="Farpl protein"
/protein_id="AAH27077.1"
/db_xref="GI:20071585"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
/translation="YKGVLRICKKHPPHNADFRCDRAALAEITKVAQLHQTMRME
NPKLHELKLDLIGDILVTPGRFRLQSLKLSKGIQOMFIFNDVLLYTSRGL
TYSNPKHVGOLPLVGMTIEESEBEWGHCLTRQROSIIVAASSREMKEMDI
QMAIDAEKNGPTELLASPPDDNSPEARADSEDDLSASTSLERAPAHGN
THVTVKHSSTVSVDSEIYAVBNQISGNLKLFRKSNQKMLVFTNFCPEFYSH
QDSHPLASLPLLGYSLTTPSESEBNHKDVPFKLHFKSHVYFRAESEYFERMEVIR
SATSSASRSLHSHKESHLV"
250..501

misc_feature /gene="Farpl1"
/note="PH: Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntrophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"
/db_xref="CCD:smarc00233"
730..1020
/gene="Farpl1"
/note="PH: Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntrophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"
/db_xref="CCD:smarc00233"

ORIGIN

Query Match 25.8%; Score 889.6; DB: 6; Length 2431;
Best Local Similarity 89.0%; Pred. No. 2,6e-241;
Matches 961; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
QY 2122 TACACGAGGCTCTGAGGCGCTGTGCAACACACCCCGAGCCAGCCGACTTCAGG 2181
DB 1 TATACGATGTCTTGAGAGGCTGTCCAGACACACCCACCAACGCGGACTTCAGG 60
QY 2182 GACTGCCAGGCGCTTTCGACAGATACGAGATGTGGCAAGCTCCAGGTTACGATG 2241
DB 61 GACTGCAGAGCTGCGCTGCGGAGATCACAGATGTGGCAAGCTGCATGTACATG 120
QY 2242 ATCAAGATGGAATTTCCAGAAAGCTGCAGAACTCAAGAAATTTGATGTGCATTCAC 2301
DB 121 ATCAAGATGGAATTTCCAGAAAGCTGCAGAACTCAAGAAATTTGATGTGCATTCAC 180
QY 2302 AATCTGTGTTCGCGGAGAGGATTCATCCGCTGTGGAGAGCTTCAGAAAGCTTCGGGG 2361
DB 181 AATCTGTGTTCGCGGAGAGGATTCATCCGCTGTGGAGAGCTTCAGAAAGCTTCGGGG 240
QY 2362 AAGGGGCTCAGAGCGATGTTCTTCGTTCAAGCAGCTCTGTATACAGAGCCGG 2421
DB 241 AAGGGGCTCAGAGCGATGTTCTTCGTTCAAGCAGCTCTGTGTATACAGAGCCGG 300
QY 2422 GGGCTACGCGCTCCCAATAGTTTAAAGTCCAGCGGAGAGCTCCGCTTATAGGCAATGAG 2481
DB 301 GAGCTGACGCACTTATAGTTTAAAGTCCAGCGGAGAGCTCCGCTTATAGGCAATGAG 360
QY 2482 ATTGAGAGAGGAGAGAGAGTGGGGGTGCCCACTGCTGACCTTCGGGGCCAGCGG 2541

[illegible]

TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS	JOURNAL	REMARK	COMMENT
<p>Worley, K.C., Hale, S.C., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodegren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Gilwood, J., Schmutz, J., Myer, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.</p> <p>Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences</p> <p>Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)</p> <p>12477932</p> <p>2 (bases 1 to 2296)</p> <p>Straubeberg, R.</p> <p>Direct Submission</p> <p>Submitted (09-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA</p> <p>NIH-MGC Project URL: http://mgc.nci.nih.gov</p> <p>Contact: MGC help desk</p> <p>Email: cgabs-r@mail.nih.gov</p> <p>Tissue Procurement: Gilbert Smith, Ph.D.</p> <p>CNA Library Preparation: Life Technologies, Inc.</p> <p>CNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)</p> <p>DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada</p> <p>info@cgsc.bc.ca</p> <p>Steven Jones, Jennifer Aasno, Ian Bosdet, Yaron Butterfield, Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Heiao, Martin Krzywnicki, Reta Kutsche, Oliver Lee, So Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeary, Steven Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeei, Jacqueline Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.</p>	<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ULNL at: http://image.lim.gov</p> <p>Series: IRAC Plate: 86 Row: O Column: 23</p> <p>This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.</p>	<p>location/Qualifiers</p> <p>1. 2296</p> <p>/organism="Mus musculus"</p> <p>/mol_type="mRNA"</p> <p>/strain="FVB/N-3"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="IMAGE:5374219"</p> <p>/tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."</p> <p>/clone_1lb="NCI CGAP_Mam2"</p> <p>/lab_host="DH10B"</p> <p>/note="Vector: PCMV-SpORT6"</p> <p><1. 2296</p> <p>/gene="Farp1"</p> <p>/note="synonym: Cdep"</p> <p>/db_xref="GeneID:223254"</p> <p>/db_xref="MGI:2446173"</p> <p><1. 926</p> <p>/gene="Farp1"</p> <p>/codon_start=3</p> <p>/product="Farp1 protein"</p> <p>/protein_id="AAH43327.1"</p> <p>/db_xref="MGI:28175182"</p> <p>/db_xref="GeneID:223254"</p> <p>/db_xref="MGI:2446173"</p> <p>/translation="KLHELEKDIIGIDNLTVPGRFIRLGSLSKLSGKGIQORMFLENDVLTASRLTASNOFKVGGPLMYGTIEESBEEGVPCHLIRGOROSIIVAASSRSMSEMRMEDIOMADILAESKNSPPELILASPPDNKSPDEATADSESDIIASRTSLEKQAPHGKNTVHYCHMKRSTSVSNVDSIAYENQNSGMLRKTKNSNMOKIWTYFVITNCLFFVYSHQDSHPLASDPLGIYSULITPESBENHKDVFVKLHFSKSHYTFRAESE</p>	<p>gene</p> <p>CDS</p>				

mlsc_feature
YPERMMEVIRSATSSASRAHLSHKESHLV
111..362

/gene="Rarp1"
/note="PH; Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntenophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinaemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol 1phide"
/db_xref="CD:smart00233"
591..881

mlsc_feature

/gene="Rarp1"
/note="PH; Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntenophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinaemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol 1phide"
/db_xref="CD:smart00233"

ORIGIN

Query Match 22.6%; Score 777.8; DB 6; Length 2296;
Best Local Similarity 89.2%; Pred. No. 1.9e-209;
Matches 839; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 2261 AGAAGCTGCGAAGCTCAAGAAAGATTGATGGCATGACATCTGTGTTCCGGGAA 2320
DB 1 AGAAGCTGCGAAGCTCAAGAAAGATTGATGGCATGACATCTGTGTTCCGGGAA 60
QY 2321 GGGAGTTTCATCGCTCTGGGAGCCTCAGCAAGCTCTCGGGGAAAGGGCTCCAGCAGGCA 2380
DB 61 GGGAGTTTCATCGCTCTGGGAGCCTCAGCAAGCTCTCGGGGAAAGGGCTCCAGCAGGCA 120
QY 2381 TGTTCCTCTGTTCAAGCAGTCTGCTATACAGAGCCGGGGCTGACGGCTCCATC 2440
DB 121 TGTTCCTCTGTTCAAGCAGTCTGCTGATACAGCCGGGGCTGACGGCTCCATC 180
QY 2441 AGTTTAAAGTCCAGGGGAGCTCCGCTATAGGCAATGAGATTGAGGAGGAAAGAGC 2500
DB 181 AGTTTAAAGTCCAGGGGAGCTCCGCTATAGGCAATGAGATTGAGGAGGAAAGAGC 240
QY 2501 AGTGGGGGAGTCCCACTGCTGACCTCCGGGGCCAGCGGCACTCATCATGTGGCCG 2560
DB 241 AGTGGGGGAGTCCCACTGCTGACCTCCGGGGCCAGCGGCACTCATCATGTGGCCG 300
QY 2561 CCAATTCTCGGTCAGAGATGAGAGAGTGGGTTAGAGCATCCAGATGGCAATTGACCTGG 2620
DB 301 CCAATTCTCGGTCAGAGATGAGAGAGTGGGTTAGAGCATCCAGATGGCAATTGACCTGG 360
QY 2621 CGGAGAAAGACACAGCCCGCCCTGAGTTCCCTGGCCAGACAGCCCGCCGACACAGT 2680
DB 361 CGGAGAAAGACACAGCCCGCCCTGAGTTCCCTGGCCAGACAGCCCGCCGACACAGT 420
QY 2681 CCCCTGATGAAGCCACCGCGGCTGACCAAGAGTCAAGAGATGACCTGAGCGCTCGGCA 2740
DB 421 CCCCTGATGAAGCCACCGCGGCTGACCAAGAGTCAAGAGATGACCTGAGCGCTCGGCA 480
QY 2741 CATGCTGAGAGCCGACGGCCCGGACCCGGGCAACACATGTTGTCAGTGTGCTGGGACC 2800
DB 481 CATGCTGAGAGCCGACGGCCCGGACCCGGGCAACACATGTTGTCAGTGTGCTGGGACC 540

QY 2801 GCACACACAGCGTCTCCATGATGAGACTTTCAGCATGCGACTGAGAGATCAGTTGCTGAA 2860
DB 541 GCACACACAGCGTCTCCATGATGAGACTTTCAGCATGCTGTGAGAGACAGCTCTCGGA 600
QY 2861 ACCTGCTGAGAAATTCAAAACAGCACGGGTGCGAGAGCTGTGGTGTGTTACAA 2920
DB 601 ACTTGTGAGAAATTCAAAACAGCACGGGTGCGAGAGCTGTGGTGTGTTACAA 660
QY 2921 ACTTGTGAGAAATTCAAAACAGCACGGGTGCGAGAGCTGTGGTGTGTTACAA 2980
DB 661 ACTTGTGAGAAATTCAAAACAGCACGGGTGCGAGAGCTGTGGTGTGTTACAA 720
QY 2981 TGCTGCTGACTGCTGCTACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTAGTGT 3040
DB 721 TGCTGCTGACTGCTGCTACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTAGTGT 780
QY 3041 TCAGCTGCACTTCAAGTCCACGCTTACTTCAAGGCGGAAAGGAGTACAGCTTGC 3100
DB 781 TCAGCTGCACTTCAAGTCCACGCTTACTTCAAGGCGGAAAGGAGTACAGCTTGC 840
QY 3101 AAGGTGATGGAAGATGCCGAGTCCAGACAGCTCTGCGGACCCCAAGCTTGA 3160
DB 841 AAGGTGATGGAAGATGCCGAGTCCAGACAGCTCTGCGGACCCCAAGCTTGA 900
QY 3161 GCCACAAAGATCTCTGTGATGATGAGCCGAGACACACTC 3201
DB 901 GTACAAAGATCTCTGTGATGATGAGCCGAGTGAACGCGCC 941

Search completed: May 22, 2006, 07:58:51
Job time : 12710.4 secs

This Page Blank (user)

Copyright (c) 1993 - 2006 Blocc

OM nucleic - nucleic search, using SW model

Run on: May 22, 2006, 00:29:52 ; Search time 1347.07 seconds
(without alignments)
17815.260 Million cell updates/sec

Title: US-09-555-342B-1

FEELER SCOR: 3442
Sequence: 1 cgccgcagcgcgcgctg.....agtatlaaacattgtcatc 3442

Scoring cable: IDENTITY_NUC

Searched: 5244920 reqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

```
Database :
1: N_GeneSeq.8.*
2: geneeqn1980s.*
3: geneeqn1990s.*
4: geneeqn2000s.*
5: geneeqn2001as.*
6: geneeqn2001bs.*
7: geneeqn2002as.*
8: geneeqn2002bs.*
9: geneeqn2003as.*
10: geneeqn2003bs.*
11: geneeqn2003cs.*
12: geneeqn2003ds.*
13: geneeqn2004as.*
14: geneeqn2004bs.*
15: geneeqn2005s.*
16: geneeqn2006s.*
```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3442	100.0	3442	2	AAX79183	Aax79183 Human cho
2	3442	100.0	3442	15	AEBO3370	AeBo3370 Acute mye
3	3438.8	99.9	5028	14	AEER27174	AeB27174 Human col
4	3437.2	99.9	4687	3	AAAO8582	AAaO8582 Human cyt
5	3170.2	92.1	3187	5	AASe64760	AAse64760 DNA encod
6	1688.6	49.1	3094	3	AAC98992	Aac98992 Human pan
7	1037.4	30.1	3997	12	ADDR4807	Adg4807 Human lunr
8	1037.4	30.1	3997	13	ADDR5675	Adr5675 Breast ca
9	1037.4	30.1	3997	7	ADRS6452	AdB6452 Human K1A
10	620.6	18.0	637	14	ACLS5653	ACl5653 Human col
11	554.2	16.1	582	10	ACD92278	ACd92278 Human col
12	554.2	16.1	582	10	ACD94832	ACd4832 Human col
13	490.8	14.3	6503	13	ACGN1157	ACn1157 Human dia
14	474.4	13.8	506	6	ABL67912	ABl67912 Ovary can
15	447	13.0	3200	12	ADOC3863	Ado3863 Novel hur
16	441	12.8	441	15	AEPR8734	AeP8734 Human CDH
17	440.8	12.8	671	15	ADGS2709	Adg2709 Novel can
18	420.8	12.2	435	14	ACCL1468	ACl1468 Human col

19	402.4	11.7	1718	4	AAS27016
20	403.4	11.7	1718	10	ADB93194
21	385.6	11.2	393	9	ACH21582
22	334.8	9.7	6565	13	ACHN1562
23	306.6	8.9	3275	13	ADRO7347
24	304.8	8.9	12277	10	ADCT71321
25	297	8.6	379	8	ABX34989
26	293.4	8.5	455	9	ACH47450
27	292	8.5	455	9	ACH4951
28	254	7.4	1824	8	AAS64761
29	229	6.7	1220	4	ABL02335
30	223.6	6.5	241	3	AAA41150
31	207.8	6.0	716	4	AAS27440
32	207.8	6.0	716	10	ADB93618
33	202.8	5.9	1747	4	ABL03867
34	200.6	5.8	710	6	ABL90828
35	193.4	5.6	2410	8	ACC44336
36	193.4	5.6	2696	14	AEAL9659
37	193.4	5.6	2867	12	ADP10315
38	192.4	5.6	2867	13	ADR30365
39	192.4	5.6	2867	13	ADP54688
40	192.4	5.6	2867	14	ADY15865
41	193.4	5.6	4938	14	AEAL9657
42	193.4	5.6	4322	5	ABV30214
43	187	5.4	2882	14	ADZ62844
44	181.8	5.3	2781	14	ADV43235
45	181.8	5.3	3643	6	ABK84074

ALIGNMENTS

ULT 1
79183

AAx79183 standard; DNA; 3442 BP

AAX79183;

17-AUG-1999 (first entry)

Human chondrocyte-derived gene CDEP.

Differentiation; human; foetal; chondrocyte; ezrin-like domain; cancer;

drug; ss.

Homo sapiens.

W09928458-A1.

10-JUN-1999

27-NOV-1998; 98WO-JP005348.

27-NOV-1997; 97JP-00342060.

(CHUS) CHUGAI SEIYAKU KK.

Kato Y, Kawamoto T, Koyano Y;

WPI; 1999-371117/31.
D. DeDD: 2007403

UNITED STATES

it.

Claim 5; Fig 1; 59pp; Japanese.

This sequence represents the coding region for a protein (CDEP) expressed in differentiated human foetal chondrocytes which contains an ezrin-like

domain, a Dbl homology (DH) domain and a pleckstrin homology (PH) domain. The nucleic acid or protein can be used in the investigation and

treatment of cancers and arthritic diseases (including chronic rheumatoid

CC arithmetic), or for screening of candidate anticancer drugs
XX Sequence 3442 BP; 864 A; 952 C; 927 G; 639 T; 0 U; 0 Other;

Query Match 100.0%; Score 3442; DB 2; Length 3442;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 3442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CCCCCAGCCGCGCGCTGTGAGATATCTCTAAGCCGCTTTTCATCATGAGAGAAA 60
DB 1 CCCCCAGCCGCGCGCTGTGAGATATCTCTAAGCCGCTTTTCATCATGAGAGAAA 60
QY 61 GAGCAGAGCCGACCCCAAGATACAGACTGAGGAGCCCGGAAAATTGCGGATCTAGTACC 120
DB 61 GAGCAGAGCCGACCCCAAGATACAGACTGAGGAGCCCGGAAAATTGCGGATCTAGTACC 120
QY 121 TTGGAACGTGACAGAAAGCCGCCCAACACTTCAGAAAACTGCTGTCCATCAAAATC 180
DB 121 TTGGAACGTGACAGAAAGCCGCCCAACACTTCAGAAAACTGCTGTCCATCAAAATC 180
QY 181 CAGATGCTGGATGACACCCAGAGGCGATTGAAGTTCCAAAGAGCTCTGGAGAGTG 240
DB 181 CAGATGCTGGATGACACCCAGAGGCGATTGAAGTTCCAAAGAGCTCTGGAGAGTG 240
QY 241 CTGCTGGATGCAATTGCAACCACTCACTCTGAGAGGTGACTATTTTGCCCTCGAG 300
DB 241 CTGCTGGATGCAATTGCAACCACTCACTCTGAGAGGTGACTATTTTGCCCTCGAG 300
QY 301 TTTTCTGATCAAAAAAGATCAAGCTGTGGCTGATCTCTTAAACCATTGTGAACAG 360
DB 301 TTTTCTGATCAAAAAAGATCAAGCTGTGGCTGATCTCTTAAACCATTGTGAACAG 360
QY 361 ATTAGAAGGCCAAAGACGTTGTGTAGTTGTGGTGAATCTTTCCGCTGACAC 420
DB 361 ATTAGAAGGCCAAAGACGTTGTGTAGTTGTGGTGAATCTTTCCGCTGACAC 420
QY 421 ACACAACTCCAAAGAACTCAAAAGTACTGTTCGCGCTGACAGTGAAGCAGACTTG 480
DB 421 ACACAACTCCAAAGAACTCAAAAGTACTGTTCGCGCTGACAGTGAAGCAGACTTG 480
QY 481 GCTCAAGGAGGTGACGTGTATATGACACAGCGAGCTCTTGATTTTCAACATTTGG 540
DB 481 GCTCAAGGAGGTGACGTGTATATGACACAGCGAGCTCTTGATTTTCAACATTTGG 540
QY 541 CAATCGAGATTGGGGATTTTGTGTAAGCTTGAGAGCTTGACAGAGACACTTAGCAAAAATA 600
DB 541 CAATCGAGATTGGGGATTTTGTGTAAGCTTGAGAGCTTGAGAGAGCACTTAGCAAAAATA 600
QY 601 TACATACCTCAGCAAGACGCACTAGAGACAAAATCGTGAATTTTCCATCAACCACTT 660
DB 601 TACATACCTCAGCAAGACGCACTAGAGACAAAATCGTGAATTTTCCATCAACCACTT 660
QY 661 GAGCAAAACACCGAGAGATCAAGATTTCCAGCTCTTAGAGATTGCCGTGGCTAGAGATG 720
DB 661 GAGCAAAACACCGAGAGATCAAGATTTCCAGCTCTTAGAGATTGCCGTGGCTAGAGATG 720
QY 721 TATGAAATCCGGTTGCACCCGGCAAGAGACAGGAAAGGACAAAGATCAATCTGGCCGTT 780
DB 721 TATGAAATCCGGTTGCACCCGGCAAGAGACAGGAAAGGACAAAGATCAATCTGGCCGTT 780
QY 781 GCCAACACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCC 840
DB 781 GCCAACACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCC 840
QY 841 AAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGCTTTCATCAAGTCTCGGCAAGTGC 900
DB 841 AAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGCTTTCATCAAGTCTCGGCAAGTGC 900
QY 901 AATAAGCGTACAGAGATCTTGAATTCCTGATGAGCAGTCCGGATTTTCTGAAGTCC 960
DB 901 AATAAGCGTACAGAGATCTTGAATTCCTGATGAGCAGTCCGGATTTTCTGAAGTCC 960
QY 961 TTCTGAAAAATCTGTGTGAAACATCATGCTTCTTTAGACTTTTGAAGAGCCCAACCA 1020
```

```
DB 961 TTCTGAAAAATCTGTGTGAAACATCATGCTTCTTTAGACTTTTGAAGAGCCCAACCA 1020
QY 1021 AAGCCCAAGCCCGCTCTTTTAAAGCCGGGATCATATTTGCTTCAAGTGTGACTAG 1080
DB 1021 AAGCCCAAGCCCGCTCTTTTAAAGCCGGGATCATATTTGCTTCAAGTGTGACTAG 1080
QY 1081 AAGCAGTTCTGACTATGTTAAAGAAAGAGACATAAGAAAGTGCAGTTTGAAGAAG 1140
DB 1081 AAGCAGTTCTGACTATGTTAAAGAAAGAGACATAAGAAAGTGCAGTTTGAAGAAG 1140
QY 1141 CACAGCAAGATTCATTATTCGAGAGCTTGTCTTCAAGCTTACAGAACTGAATTCGAA 1200
DB 1141 CACAGCAAGATTCATTATTCGAGAGCTTGTCTTCAAGAGCTTACAGAACTGAATTCGAA 1200
QY 1201 GTGCTGGAGCAGTCTCAGCAGAGACAGACCTTACATTTGAGAAAGTGCAGAACTCCA 1260
DB 1201 GTGCTGGAGCAGTCTCAGCAGAGACAGACCTTACATTTGAGAAAGTGCAGAACTCCA 1260
QY 1261 GGGGGCCAGAGCTGCGCGGAGAAAGAACCGAAGTTTCCGCGGGGAGCCGGGGTCTG 1320
DB 1261 GGGGGCCAGAGCTGCGCGGAGAAAGAACCGAAGTTTCCGCGGGGAGCCGGGGTCTG 1320
QY 1321 CACCCGAGCCTTGGCGCGAGAGAAAGCCCGGGGTAAACAAGCAGGCGAGCGAGCCGCC 1380
DB 1321 CACCCGAGCCTTGGCGCGAGAGAAAGCCCGGGGTAAACAAGCAGGCGAGCGAGCCGCC 1380
QY 1381 TCGGGGCCACGAGAGAAAGAGAGGTGTTAAGATPAGAACCCAGACAGTAAACT 1440
DB 1381 TCGGGGCCACGAGAGAAAGAGAGGTGTTAAGATPAGAACCCAGACAGTAAACT 1440
QY 1441 CAGCCCCCGCAGCCAAAGCAGAGCTCCCTGACTGGCAGTCTCTCACTTTCCGAGCTGTCT 1500
DB 1441 CAGCCCCCGCAGCCAAAGCAGAGCTCCCTGACTGGCAGTCTCTCTCACTTTCCGAGCTGTCT 1500
QY 1501 GTGAATCTGCAAGGGGGAGTGGCCCTTGCCTCAAGCTGACTTGTCTCCAACTGAGCCCC 1560
DB 1501 GTGAATCTGCAAGGGGGAGTGGCCCTTGCCTCAAGCTGACTTGTCTCCAACTGAGCCCC 1560
QY 1561 GACACCAAGCAGGCGCTCCCTGTGATCAGCCCGGCTGAATGACAGAGCTGAGCCCCGG 1620
DB 1561 GACACCAAGCAGGCGCTCCCTGTGATCAGCCCGGCTGTGAATGACAGAGCTGAGCCCCGG 1620
QY 1621 ACGAGATGAGAGATGAGGGCCGAGAGAGATTTCCCACTGATTAAGCGTACTTCATA 1680
DB 1621 ACGAGATGAGAGATGAGGGCCGAGAGAGATTTCCCACTGATTAAGCGTACTTCATA 1680
QY 1681 GCTAAGAAAGTGTCTTACACCGAGAGACATATCTGAAGGATCTCGAAGTTATCACTTCG 1740
DB 1681 GCTAAGAAAGTGTCTTACACCGAGAGACATATCTGAAGGATCTCGAAGTTATCACTTCG 1740
QY 1741 TGGTTTCAGAGACAGTGAAGAGAGAGCGCATGCGGAGAGACTGAAGAAAGTCTCAT 1800
DB 1741 TGGTTTCAGAGACAGTGAAGAGAGAGCGCATGCGGAGAGACTGAAGAAAGTCTCAT 1800
QY 1801 TTCCCGAATTTTGAACCTTTGACAAATTTTCACTAATTTTCTCAAGAAATTGAGCAA 1860
DB 1801 TTCCCGAATTTTGAACCTTTGACAAATTTTCACTAATTTTCTCAAGAAATTGAGCAA 1860
QY 1861 GCACTTGCCTCTGTGGAAAGGCGCTCAAAATGCCCAATCAAGATTTACCAAGAAATCGGC 1920
DB 1861 GCACTTGCCTCTGTGGAAAGGCGCTCAAAATGCCCAATCAAGATTTACCAAGAAATCGGC 1920
QY 1921 GATGTCATGCTGAAGAACATTCAAGGACGTGAAGCCTGAGGAGCTCACTGTGAAGAGC 1980
DB 1921 GATGTCATGCTGAAGAACATTCAAGGACGTGAAGCCTGAGGAGCTCACTGTGAAGAGC 1980
QY 1981 AGCAGAGCCTTGAAGGCCCTGAGAAATGAATCAAGAGCTCCCGGCGCTGAGAACTTC 2040
DB 1981 AGCAGAGCCTTGAAGGCCCTGAGAAATGAATCAAGAGCTCCCGGCGCTGAGAACTTC 2040
QY 2041 TGCAGAGACTTTGAGCTGAGAGAGTGTGTTAACCTACGCTCAACACTTCTCTGCGGG 2100
```

```
Db 2041 TGCAGAGACTTTAGCTGCGAAGAGTGTGTTACTTACCGCTCAACACTTCTCTGCGG 2100
Qy 2101 CCACTGCAACGGGTGATGCACTACAGCAGAGTCTTGAGAGGGGTGTGCAAAACCAACCGG 2160
Db 2101 CCACTGCAACGGGTGATGCACTACAGCAGAGTCTTGAGAGGGGTGTGCAAAACCAACCGG 2160
Qy 2161 CCGAGCAGCGGCACTTCAAGGAGCTGCGGAGCGCTTTGGCAGAGATCAACGAGATGATG 2220
Db 2161 CCGAGCAGCGGCACTTCAAGGAGCTGCGGAGCGCTTTGGCAGAGATCAACGAGATGATG 2220
Qy 2221 GCAACAGCTCAACGGTACGATGATCAAGATGAGAAATTTCCAGAGCTGCAACGACTCAAG 2280
Db 2221 GCAACAGCTCAACGGTACGATGATCAAGATGAGAAATTTCCAGAGCTGCAACGACTCAAG 2280
Qy 2281 AAAAGATTTGATTTGGCATTTGATCAATCTTGTGTGTTCCGGGAAAGGAGTTCACTCGTGGG 2340
Db 2281 AAAAGATTTGATTTGGCATTTGATCAATCTTGTGTGTTCCGGGAAAGGAGTTCACTCGTGGG 2340
Qy 2341 AGCTTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAGGCAATGTTCTTCCGTTCAACGAC 2400
Db 2341 AGCTTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAGGCAATGTTCTTCCGTTCAACGAC 2400
Qy 2401 GTCTGCTATATACAGAGCCGGGGCTGACGGCTCCCAATCAGTTTAAAGTCCAGGGCAG 2460
Db 2401 GTCTGCTATATACAGAGCCGGGGCTGACGGCTCCCAATCAGTTTAAAGTCCAGGGCAG 2460
Qy 2461 CTCCCGCTCTATGCGCATGACGATTTGAGAGAGCGAAAGCAGTGGGGGGTGCCTCCACTGC 2520
Db 2461 CTCCCGCTCTATGCGCATGACGATTTGAGAGAGCGAAAGCAGTGGGGGGTGCCTCCACTGC 2520
Qy 2521 CTGACCTCTCGGGGGCCAGGGGCGAGTCAATCATCTGTGGCGGCACTCTGCGTCCGAGATG 2580
Db 2521 CTGACCTCTCGGGGGCCAGGGGCGAGTCAATCATCTGTGGCGGCACTCTGCGTCCGAGATG 2580
Qy 2581 GAGAACTGGGTTGAGGACATCCAGATGCGCAATTGACCTGCGGAGAAAGACAGCAGCCCC 2640
Db 2581 GAGAACTGGGTTGAGGACATCCAGATGCGCAATTGACCTGCGGAGAAAGACAGCAGCCCC 2640
Qy 2641 GCGCCCTGAGTTCTTGCGCCAGCAGCCCCCTTGACAAACAGTCCCTATGAAAGCAGCCGG 2700
Db 2641 GCGCCCTGAGTTCTTGCGCCAGCAGCCCCCTTGACAAACAGTCCCTATGAAAGCAGCCGG 2700
Qy 2701 GCTGACCAAGAGTCAGAGATGACCTGAGGCGCTCGCGCACTCGCTGAGCGCCAGGCC 2760
Db 2701 GCTGACCAAGAGTCAGAGATGACCTGAGGCGCTCGCGCACTCGCTGAGCGCCAGGCC 2760
Qy 2761 CCGCACCGCGGCAACCAATGTGTGCACTGTGTGCGCACCGCAACCAACGCTCTCCATG 2820
Db 2761 CCGCACCGCGGCAACCAATGTGTGCACTGTGTGCGCACCGCAACCAACGCTCTCCATG 2820
Qy 2821 GTGGAATTCAGATTCGCAAGTGAATCACTGTGTGGAATCCTGCTGAGAAATTCAA 2880
Db 2821 GTGGAATTCAGATTCGCAAGTGAATCACTGTGTGGAATCCTGCTGAGAAATTCAA 2880
Qy 2881 GTGGAATTCAGATTCGCAAGTGAATCACTGTGTGGAATCCTGCTGAGAAATTCAA 2940
Db 2881 GTGGAATTCAGATTCGCAAGTGAATCACTGTGTGGAATCCTGCTGAGAAATTCAA 2940
Qy 2941 AAATGACACACAGGACAAATCATCCCTTGGCAGCTCTCTGCTGCGCTACTGCTCAAC 3000
Db 2941 AAATGACACACAGGACAAATCATCCCTTGGCAGCTCTCTGCTGCGCTACTGCTCAAC 3000
Qy 3001 ATCCCTCTGAGTTCGCAAGATTCGCAAGTGAATCACTGTGTGGAATCCTGCTGAG 3060
Db 3001 ATCCCTCTGAGTTCGCAAGATTCGCAAGTGAATCACTGTGTGGAATCCTGCTGAG 3060
Qy 3061 CACGTTACTACTTCAAGGCGGAAAGCGATACAGCTTCAAGAGTGTGATGAGATGATC 3120
Db 3061 CACGTTACTACTTCAAGGCGGAAAGCGATACAGCTTCAAGAGTGTGATGAGATGATC 3120
Qy 3121 CCGAGTGCACCAAGCTCTGCTGCGACCCCAAGTGTGAGCCCAAGAGTCTTGTG 3180
Db 3121 CCGAGTGCACCAAGCTCTGCTGCGACCCCAAGTGTGAGCCCAAGAGTCTTGTG 3180
```

```
Qy 3181 TATTGATGGCGGACACACTGCTTCCGAGTGGGCTTTCTCTGSAAGAGCTTCTT 3240
Db 3181 TATTGATGGCGGACACACTGCTTCCGAGTGGGCTTTCTCTGSAAGAGCTTCTT 3240
Qy 3241 CTTCGTATTAATGAAAGCTGTGTAATTAACACCTGTCTGAAATCAAAAAATGAGCTT 3300
Db 3241 CTTCGTATTAATGAAAGCTGTGTAATTAACACCTGTCTGAAATCAAAAAATGAGCTT 3300
Qy 3301 CCGAGCAGCTCTCTGCTCTGCAACAGCGCGCTTTTAAACCCGACCTCTCAGCTTTGAA 3360
Db 3301 CCGAGCAGCTCTCTGCTCTGCAACAGCGCGCTTTTAAACCCGACCTCTCAGCTTTGAA 3360
Qy 3361 TGAACAGCGCTCCACCTCAGTCCGATCCGATCCGATCCGATCCGATCCGATCCGATCC 3420
Db 3361 TGAACAGCGCTCCACCTCAGTCCGATCCGATCCGATCCGATCCGATCCGATCCGATCC 3420
Qy 3421 CCAATATTAATTAATTTGATTT 3442
Db 3421 CCAATATTAATTAATTTGATTT 3442
```

```
RESULT 2
AEB03370
ID AEB03370 standard; DNA; 3442 BP.
XX
AC AEB03370;
XX
DT 09-FEB-2006 (first entry)
XX
DE Acute myelogenous leukemia prognosis related DNA sequence SEQ ID NO: 31.
XX
KW ds; gene; acute myelogenous leukemia; prognosis; gene expression;
XX bioclip.
XX
OS Homo sapiens.
XX
PN JP200533987-A.
XX
PD 08-DEC-2005.
XX
PF 06-MAY-2005; 2005JP-00135284.
XX
PR 06-MAY-2004; 2004US-0566635P.
XX
PA (VERI-) VERIDEX LLC.
XX
PI Laponi M;
XX
PS WPI; 2006-004067/01.
XX
CC The present sequence is that of a human gene sequence which is claimed
CC for use in evaluating the prognosis of patients suffering from acute
CC myelogenous leukemia (AML) by analysis of the patients gene expression
CC profile. The invention relates to a novel method for evaluating the
CC prognosis of a patient with acute myelogenous leukemia by detecting
CC higher/lower expression level of genes encoding mRNA recognized by a
CC probe set chosen from 19 nucleotide sequences (AEB03340 or AEB03344-
CC AEB03361). Also, gene expression profiling of AML patients using a set of
CC probes based on 167 sequences (AEB03362-AEB03527) using a bioclip was
CC performed to determine prognosis. Methods are also included for
CC determining an AML patient's state, determining AML patient's treatment
CC protocol, determining whether the patient will respond to the treatment,
CC and producing an AML patient's prognosis report by analysis of the
CC patients gene expression profile. The methods are useful in evaluating an
CC AML patient's prognosis and for treating AML patient.
```

XX Sequence 3442 BP; 864 A; 952 C; 927 G; 699 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 3442; DB 15; Length 3442;
Beet Local Similarity 100.0%; Pred. No. 0; Mismatches 3442; Conservative 0; Indels 0; Gaps 0;
Matches 3442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CCCCCAGCCGCGCGCTGTGGAGATATCTTAAGCCGCTTTCATCATGGAGAAATA 60
Db 1 CCCCCAGCCGCGCGCTGTGGAGATATCTTAAGCCGCTTTCATCATGGAGAAATA 60
QY 61 GAGCAGAGGCCGACCCCGAGATACAGACTGGGGGCCCCGAAAATTCCGGGATCACTACC 120
Db 61 GAGCAGAGGCCGACCCCGAGATACAGACTGGGGGCCCCGAAAATTCCGGGATCACTACC 120
QY 121 TTGGAACGTGSA CAGAAAGCCGCCCCCAACCTTCACGAAA CTGCTGTCATCAAAAATC 180
Db 121 TTGGAACGTGSA CAGAAAGCCGCCCCCAACCTTCACGAAA CTGCTGTCATCAAAAATC 180
QY 181 CAGATGCTGGATGACACCCAGAGAGCATTTGAAGTTCCACAAAAGACTTCCTGGGAAGTG 240
Db 181 CAGATGCTGGATGACACCCAGAGAGCATTTGAAGTTCCACAAAAGACTTCCTGGGAAGTG 240
QY 241 CTGCTGGATGCA GTTTGCAACCACTCAACCTCTGGAAGGTGACTATTTGGCCTCGAG 300
Db 241 CTGCTGGATGCA GTTTGCAACCACTCAACCTCTGGAAGGTGACTATTTGGCCTCGAG 300
QY 301 TTTCCTGATCAAAAAGATCA CGGTGTGGCTGGAATCTCTAAAACCCATTTGAAAACG 360
Db 301 TTTCCTGATCAAAAAGATCA CGGTGTGGCTGGAATCTCTAAAACCCATTTGAAAACG 360
QY 361 ATTAGAAGCCAAAGCACGTTGTTTAAAGTTTGTTGTAATAATCTTTCCGCTGACAC 420
Db 361 ATTAGAAGCCAAAGCACGTTGTTTAAAGTTTGTTGTAATAATCTTTCCGCTGACAC 420
QY 421 ACACACTCTCAAGAAATCTCACAGGTACCTGTTGCGCTGAGGTGAAGCAGACTTG 480
Db 421 ACACACTCTCAAGAAATCTCACAGGTACCTGTTGCGCTGAGGTGAAGCAGACTTG 480
QY 481 GGTCAAGGAGGTGACGTGTAATGACACAGGCGACTCTGTGATTTCAACATTTGG 540
Db 481 GGTCAAGGAGGTGACGTGTAATGACACAGGCGACTCTGTGATTTCAACATTTGG 540
QY 541 CAATCTGAGATTTGGGATTTTATGTAAGCTTGACAGAGACACTTAAACAAAATAAA 600
Db 541 CAATCTGAGATTTGGGATTTTATGTAAGCTTGACAGAGACACTTAAACAAAATAAA 600
QY 601 TACATTAATCTCAGCAAGACGCACTAAGAGCAAAATCTGGAATTTCAACATTAACCAATT 660
Db 601 TACATTAATCTCAGCAAGACGCACTAAGAGCAAAATCTGGAATTTCAACATTAACCAATT 660
QY 661 GGACAAAACACGAGCAATCAGATTTCCAGCTCCTAAGAGATTGCCGTGGCTAGAGATG 720
Db 661 GGACAAAACACGAGCAATCAGATTTCCAGCTCCTAAGAGATTGCCGTGGCTAGAGATG 720
QY 721 TATGGAATCCGGTTGACCCGCGCAAGGACAGGAAAGGCAAGAAATCAATCTGGCCGTT 780
Db 721 TATGGAATCCGGTTGACCCGCGCAAGGACAGGAAAGGCAAGAAATCAATCTGGCCGTT 780
QY 781 GCCAACACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGGC 840
Db 781 GCCAACACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGGC 840
QY 841 AAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGTTTCTCATCAAGCTCCGGCCAGATGCC 900
Db 841 AAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGTTTCTCATCAAGCTCCGGCCAGATGCC 900
QY 901 AATAGTGCCTTACAGAGATCTTGAATTCCTGATGCGCAGTCGGGATTTCTGCAAGTCC 960
Db 901 AATAGTGCCTTACAGAGATCTTGAATTCCTGATGCGCAGTCGGGATTTCTGCAAGTCC 960
QY 961 TTCTGGAAAATCTGTGTTGAACATCATGCTTCTTTAGACTTTTGAAGGCCCAACCA 1020

Db 961 TTCTGGAAAATCTGTGTTGAACATCATGCTTCTTTAGACTTTTGAAGGCCCAACCA 1020
QY 1021 AAGCCCAAGCCGCTCCTTTAGCCGGGGGTATCATCTTTGGTTGATGTCGACTGAG 1080
Db 1021 AAGCCCAAGCCGCTCCTTTAGCCGGGGGTATCATCTTTGGTTGATGTCGACTGAG 1080
QY 1081 AAGCAGTTCTGACTATGTTAAAGAGAGACATTAAGAGTGCAGTTTGAAGAAAG 1140
Db 1081 AAGCAGTTCTGACTATGTTAAAGAGAGACATTAAGAGTGCAGTTTGAAGAAAG 1140
QY 1141 CACAGCAAGATTCATTCATTCGAGGCTTGTACAGCCCTACAGAACTGAATTCGGA 1200
Db 1141 CACAGCAAGATTCATTCATTCGAGGCTTGTGTACAGCTTACAGAACTGAATTCGGA 1200
QY 1201 GTGCTGAGACAGTCTCAGCAGAGCACAGCTTACATTTGGAGAGTGCAGAACTCCA 1260
Db 1201 GTGCTGAGACAGTCTCAGCAGAGCACAGCTTACATTTGGAGAGTGCAGAACTCCA 1260
QY 1261 GGGGGCCAGAGCTGCGCGAGAGAAAGAACCGAAGTTTCGCGCGGGAGCCGGGGTGG 1320
Db 1261 GGGGGCCAGAGCTGCGCGAGAGAAAGAACCGAAGTTTCGCGCGGGAGCCGGGGTGG 1320
QY 1321 CACCCGAGCCCTCTGCGCCGAGAGAAAGCCCGCGGTAA CAAAGAGCGGAGCGGACCGGC 1380
Db 1321 CACCCGAGCCCTCTGCGCCGAGAGAAAGCCCGCGGTAA CAAAGAGCGGAGCGGACCGGC 1380
QY 1381 TCGGCGCCACGAGAGAGAGAGAGGTCTTAAAGATAGGACCCAGACAGATAACT 1440
Db 1381 TCGGCGCCACGAGAGAGAGAGAGGTCTTAAAGATAGGACCCAGACAGATAACT 1440
QY 1441 CAGCCCCGACAGCCAAAGCA CAGGCTCCCTGACTGGCACTGCTCACCTTTCGAGCTGCT 1500
Db 1441 CAGCCCCGACAGCCAAAGCA CAGGCTCCCTGACTGGCACTGCTCACCTTTCGAGCTGCT 1500
QY 1501 GTGAACCTGCGAGGGGGAGTGCCCTGSCAACGTAACGTAACCTTGACCCGACCC 1560
Db 1501 GTGAACCTGCGAGGGGGAGTGCCCTGSCAACGTAACGTAACCTTGACCCGACCC 1560
QY 1561 GACACCAAGCAGGCTCTCCCTTGATCAGCCCGCTGTGAATGACAGAGCCTGCCCGGG 1620
Db 1561 GACACCAAGCAGGCTCTCCCTTGATCAGCCCGCTGTGAATGACAGAGCCTGCCCGGG 1620
QY 1621 ACGACGATGAGATGAGAGGCGCGAGAGAAAGATTTCCCACTGATTAAGCTTACATA 1680
Db 1621 ACGACGATGAGATGAGAGGCGCGAGAGAAAGATTTCCCACTGATTAAGCTTACATA 1680
QY 1681 GCTAAGAAAGTGTCTAACACCGAGCGAACATATCTGAAGATCTCGAAGTATCACTTGG 1740
Db 1681 GCTAAGAAAGTGTCTAACACCGAGCGAACATATCTGAAGATCTCGAAGTATCACTTGG 1740
QY 1741 TGGTTTCAGAGCACAGTAGCAAAAGAGAGCCATGCGGAAAGCACTGAAAAGTCTCAT 1800
Db 1741 TGGTTTCAGAGCACAGTAGCAAAAGAGAGCCATGCGGAAAGCACTGAAAAGTCTCAT 1800
QY 1801 TTCCGGAATTTTGAACCTTTGCACAAATTTCACTAATTTTCTCAAGAAATTAAGCA 1860
Db 1801 TTCCGGAATTTTGAACCTTTGCACAAATTTCACTAATTTTCTCAAGAAATTAAGCA 1860
QY 1861 CGACTTGGCCCTGTGGAGAGCGCTCAAAATGCCAAATCAAGAGTTTCAAAAGAAATCGGC 1920
Db 1861 CGACTTGGCCCTGTGGAGAGCGCTCAAAATGCCAAATCAAGAGTTTCAAAAGAAATCGGC 1920
QY 1921 GATGTCATGCTGAAGAACATTCAAGGCAATGAAGCACTGCGGCTCACTGTGGAAGCAC 1980
Db 1921 GATGTCATGCTGAAGAACATTCAAGGCAATGAAGCACTGCGGCTCACTGTGGAAGCAC 1980
QY 1981 AAGCAGGCTTTGAGAGCCCTGGAAGATGAATCAAGAGACTCCCGCGGCTGGAAGAACTTC 2040
Db 1981 AAGCAGGCTTTGAGAGCCCTGGAAGATGAATCAAGAGACTCCCGCGGCTGGAAGAACTTC 2040
QY 2041 TGCAGAGACTTGAAGCTGCAAGAGGTGTTAACCTACCGCTCAACACTTCTCTCGCGG 2100
Db 2041 TGCAGAGACTTGAAGCTGCAAGAGGTGTTAACCTACCGCTCAACACTTCTCTCGCGG 2100

QY 2101 CCACTGACCGGCTCATGACCTACAGAGCTCTGGAGGCGCTGCAAAACACACCCG 2160
 DB 2101 CCACTGACCGGCTCATGACCTACAGAGCTCTGGAGGCGCTGCAAAACACACCCG 2160
 QY 2161 CCGAGCAGCGCGCACTTCAGAGGACCTGCGAGCGCTTTGGCAGAGATCAACGAGATGATG 2220
 DB 2161 CCGAGCAGCGCGCACTTCAGAGGACCTGCGAGCGCTTTGGCAGAGATCAACGAGATGATG 2220
 QY 2221 GCACAGCTCCAGGATCGATGATCAAGATGAGAAATTTCCAGAGCTGCAACGACTCAAG 2280
 DB 2221 GCACAGCTCCAGGATCGATGATCAAGATGAGAAATTTCCAGAGCTGCAACGACTCAAG 2280
 QY 2281 AAGAGATTGATGAGATGATCACTTTGGGTTCCGGGAAAGGAGTTCAATCCGCTGGGAC 2340
 DB 2281 AAGAGATTGATGAGATGATCACTTTGGGTTCCGGGAAAGGAGTTCAATCCGCTGGGAC 2340
 QY 2341 AGCCTCAGCAGCTCTCGGGGAAAGGGGCTCAGCAGCGCATGTTCTTCTGTTCAACGAC 2400
 DB 2341 AGCCTCAGCAGCTCTCGGGGAAAGGGGCTCAGCAGCGCATGTTCTTCTGTTCAACGAC 2400
 QY 2401 GTCTGCTATACAGAGCCGGGGGCTGACGGGCTTCAATCAATTTAAAGTCCAGGGCAG 2460
 DB 2401 GTCTGCTATACAGAGCCGGGGGCTGACGGGCTTCAATCAATTTAAAGTCCAGGGCAG 2460
 QY 2461 CTCCCGCTTATGAGCATGACATTGAGAGAGCGAAAGACGATGGGGGGTCCCCCACTGC 2520
 DB 2461 CTCCCGCTTATGAGCATGACATTGAGAGAGCGAAAGACGATGGGGGGTCCCCCACTGC 2520
 QY 2521 CTGACCTCTCGGGGCGCAGCGGCGAGTCCATCATCTGAGGCGCGCATCTGCTGCGAGATG 2580
 DB 2521 CTGACCTCTCGGGGCGCAGCGGCGAGTCCATCATCTGAGGCGCGCATCTGCTGCGAGATG 2580
 QY 2581 GAGAGTGGGTTGAGAGCATCCAGATGCGCATTTGACCTGCGGAGAAAGACAGCAGCCCC 2640
 DB 2581 GAGAGTGGGTTGAGAGCATCCAGATGCGCATTTGACCTGCGGAGAAAGACAGCAGCCCC 2640
 QY 2641 GCGCCCTGAATTTCTGCGCCAGCAGCCCCCTTGAACAAGTCCCTATATAGAACACCGCG 2700
 DB 2641 GCGCCCTGAATTTCTGCGCCAGCAGCCCCCTTGAACAAGTCCCTATATAGAACACCGCG 2700
 QY 2701 GGTGACACAGAGTCAAGAGATGACCTGAGCGGCTTCCGCAATCGTGGAGCGCCAGGCC 2760
 DB 2701 GGTGACACAGAGTCAAGAGATGACCTGAGCGGCTTCCGCAATCGTGGAGCGCCAGGCC 2760
 QY 2761 CCGCAGCGCGGCAACAATGATGACGATGCTGCGCAGCAGCAACAGCGCTTCCATG 2820
 DB 2761 CCGCAGCGCGGCAACAATGATGACGATGCTGCGCAGCAGCAACAGCGCTTCCATG 2820
 QY 2821 GTGGAATTGAGCATCGCAGTGGAGATCACTTTGTTGAGAAACCTGCTGAGAAATTCAAA 2880
 DB 2821 GTGGAATTGAGCATCGCAGTGGAGATCACTTTGTTGAGAAACCTGCTGAGAAATTCAAA 2880
 QY 2881 AACAGCAACGGGTCGAGAGCTGAGGATGTTTCAAACTTTCGCTGTTCTTCTAC 2940
 DB 2881 AACAGCAACGGGTCGAGAGCTGAGGATGTTTCAAACTTTCGCTGTTCTTCTAC 2940
 QY 2941 AAATCAACACAGGACATATCCCTTGCAGGCTCTCTGCTGCGCTTACTGCTAC 3000
 DB 2941 AAATCAACACAGGACATATCCCTTGCAGGCTCTCTGCTGCGCTTACTGCTAC 3000
 QY 3001 ATCCCTCTGAGTCCGAGAACTACCAAGAAAGATGTTTCAAGTCTTCAAGTCTC 3060
 DB 3001 ATCCCTCTGAGTCCGAGAACTACCAAGAAAGATGTTTCAAGTCTTCAAGTCTC 3060
 QY 3061 CACGTCTACTACTTCAGGCGGAAACGAGTACACCTTCAAAAGTGAAGTATC 3120
 DB 3061 CACGTCTACTACTTCAGGCGGAAACGAGTACACCTTCAAAAGTGAAGTATC 3120
 QY 3121 CGCAGTCCACACGCTCTGCTGCGACCCACAGTGTGAGCCACAAAGATCTTTG 3180
 DB 3121 CGCAGTCCACACGCTCTGCTGCGACCCACAGTGTGAGCCACAAAGATCTTTG 3180

QY 3181 TATTGATGGCGGACACACTGTTCCGAGTGGGCTTTCTGGAAGACGTTCTTT 3240
 DB 3181 TATTGATGGCGGACACACTGTTCCGAGTGGGCTTTCTGGAAGACGTTCTTT 3240
 QY 3241 CTTCGTATTAATGAAAGCTGTTAAATTAACACCTGTCTGAAATCAAAAATGAGCTT 3300
 DB 3241 CTTCGTATTAATGAAAGCTGTTAAATTAACACCTGTCTGAAATCAAAAATGAGCTT 3300
 QY 3301 CCCAGCAGCTCTCTGCTCTCCACAGCGGCTTTTAAACCCGACCTCTCAGCGTTGAA 3360
 DB 3301 CCCAGCAGCTCTCTGCTCTCCACAGCGGCTTTTAAACCCGACCTCTCAGCGTTGAA 3360
 QY 3361 TGAACAGCGCTCCACCTCAGCTGCGATCCGCTGAGGGGCGCTGTTAGCTAGTG 3420
 DB 3361 TGAACAGCGCTCCACCTCAGCTGCGATCCGCTGAGGGGCGCTGTTAGCTAGTG 3420
 QY 3421 CCAATTTAAACATTTGTCATT 3442
 DB 3421 CCAATTTAAACATTTGTCATT 3442

RESULT 3
 AEB22714
 ID AEB22714 standard; DNA; 5028 BP.
 XX

AC AEB22714;

DT 22-SEP-2005 (first entry)

DE Human colon cancer-associated gene SEQ ID NO: 9.

XX cancer; genetic marker; lymphoma; cytostatic; hematological disease;

KW immune disorder; neoplasm; colon tumor; gastrointestinal disease;

KW adenocarcinoma; carcinoma; screening; diagnostic; prognostic;

XX pharmaceutical; ds; gene.

OS Homo sapiens.

XX WO2005064009-A1.

PD 14-JUL-2005.

PF 23-DEC-2004; 2004WO-DK000914.

XX 27-DEC-2003; 2003DK-00001940.

PR 24-JAN-2004; 2004DK-00000096.

PR 07-APR-2004; 2004DK-00000586.

PR 26-NOV-2004; 2004DK-00001843.

XX (AROS-) AROS APPLIED BIOTECHNOLOGY APS.

XX Ornofc TF, Jensen JL, Kruhoffer M, Laiho P, Aaltonen LA;

PI WPI; 2005-506661/51.

XX Claim 9; SEQ ID NO 9; 338bp; English.

XX The invention relates to a method of classifying the cancer in an

CC individual having contracted cancer, which comprises determining the

CC microsatellite status of the tumor. The microsatellite status or the

CC hereditary or sporadic nature of the cancer is determined by a method

CC comprising: (a) obtaining a sample from the individual having contracted

CC cancer, the sample comprising gene expression products; the presence

CC and/or amount of which forms a pattern that is indicative of the

CC microsatellite status or the hereditary or sporadic nature of the cancer;

CC (b) determining the presence and/or amount of the gene expression

CC products forming the pattern, and obtaining an indication of the cancer

CC microsatellite status or the hereditary or sporadic nature of the cancer

CC in the individual based on step (b). At least one or two of the gene

CC expression products forming a pattern used to determine the


```

OY 1801 TTCCCGAATTTTGAACCTTTGCACAAATTTTCTACTTAATTTTCTCAAGGAAATTTGACAA 1860
DB 2089 TTCCCGAATTTTGAACCTTTGCACAAATTTTCTACTTAATTTTCTCAAGGAAATTTGACAA 2148
OY 1861 CGACTTGGCCCTGTGGAAAGGCGCTCAAAATGCCAAATGAGAAATTTACCAAAGAAATCGGCG 1920
DB 2149 CGACTTGGCCCTGTGGAAAGGCGCTCAAAATGCCAAATGAGAAATTTACCAAAGAAATCGGCG 2208
OY 1921 GATGTATGTCTGAAGAAATTTCAAGGAGTGAAGCACTTGGCGGCTCACTGTGTGAAGCAC 1980
DB 2209 GATGTATGTCTGAAGAAATTTCAAGGAGTGAAGCACTTGGCGGCTCACTGTGTGAAGCAC 2268
OY 1981 AGGAGGAGGCGCTGGAGGCGCTGGAGAAATGGAATCAAGAGCTCCGCGGCGCTGGAGAACTTC 2040
DB 2269 AGGAGGAGGCGCTGGAGGCGCTGGAGAAATGGAATCAAGAGCTCCGCGGCGCTGGAGAACTTC 2328
OY 2041 TGCAGAGACTTTGAGCTGAGAGAAAGTGTGTACTACCTGCTCAACACTTCTCTCTGCGG 2100
DB 2329 TGCAGAGACTTTGAGCTGAGAGAAAGTGTGTACTACCTGCTCAACACTTCTCTCTGCGG 2388
OY 2101 CCACTGCAACCGGCTCATGCACTAACAAGCAGGCTCTGAGAGCGCTGTGCAAAACAACCCG 2160
DB 2389 CCACTGCAACCGGCTCATGCACTAACAAGCAGGCTCTGAGAGCGCTGTGCAAAACAACCCG 2448
OY 2161 CGAGGCAAGCGGCTTCAAGGAGCTGCGAGCGGCTTTGGCAGAGATCAACGAGAGTGTG 2220
DB 2449 CGAGGCAAGCGGCTTCAAGGAGCTGCGAGCGGCTTTGGCAGAGATCAACGAGAGTGTG 2508
OY 2221 GCACAGCTCCACGCTGATGATCAAGATGAGAAATTTTCAGAAAGCTGCAAGAACTCAAG 2280
DB 2509 GCACAGCTCCACGCTGATGATCAAGATGAGAAATTTTCAGAAAGCTGCAAGAACTCAAG 2568
OY 2281 AAAGATTTGATTTGGCATTTGACAACTTTGTGTGTCCGGGAAAGGAGTTCACTCGCTGCGC 2340
DB 2569 AAAGATTTGATTTGGCATTTGACAACTTTGTGTGTCCGGGAAAGGAGTTCACTCGCTGCGC 2628
OY 2341 AGCTTCAGAGAGCTCCGGGGGAAAGGGGCTCCAGCAGAGGATGTTCTTCCGTTTCAACGAC 2400
DB 2629 AGCTTCAGAGAGCTCCGGGGGAAAGGGGCTCCAGCAGAGGATGTTCTTCCGTTTCAACGAC 2688
OY 2401 GTCTGTCTATACAGAGCGGCGGCTGACGCGCTCCAAATCAAGTTTAAAGTCCACGGGCA 2460
DB 2689 GTCTGTCTATACAGAGCGGCGGCTGACGCGCTCCAAATCAAGTTTAAAGTCCACGGGCA 2748
OY 2461 CTCCCGCTCTATGCGCATGACGATTTGAGAGAGCGAAGAGAGTGGGGGCTGCCCACTGC 2520
DB 2749 CTCCCGCTCTATGCGCATGACGATTTGAGAGAGCGAAGAGAGTGGGGGCTGCCCACTGC 2808
OY 2521 CTGACCTCCGGGGGCGAGCGGCGATCATCTGTGTGCGGCGGCTTCTGCTCCGAGATG 2580
DB 2809 CTGACCTCCGGGGGCGAGCGGCGATCATCTGTGTGCGGCGGCTTCTGCTCCGAGATG 2868
OY 2581 GAGAAAGTGGGTTGAGAGACATCCAGATGCGCACTTGAACCTGCGGAGAAAGACAGCAGCCCC 2640
DB 2869 GAGAAAGTGGGTTGAGAGACATCCAGATGCGCACTTGAACCTGCGGAGAAAGACAGCAGCCCC 2928
OY 2641 GCCCTGAGTTCTTGCGCCAGAGCGCCCTGTACACAAGTCCCTGATGAAGCCACCGCG 2700
DB 2929 GCCCTGAGTTCTTGCGCCAGAGCGCCCTGTACACAAGTCCCTGATGAAGCCACCGCG 2988
OY 2701 GCTGACCAAGAGTCAAGAGATGACCTGAGCGGCTCGCGCACTGCTGGAAGCGCCAGGCG 2760
DB 2989 GCTGACCAAGAGTCAAGAGATGACCTGAGCGGCTCGCGCACTGCTGGAAGCGCCAGGCG 3048
OY 2761 CCGCACCGCGGCAACAATGTGTGCACTGTGTGCTGCGCACCGCAACAACAGCGCTCTCATG 2820
DB 3049 CCGCACCGCGGCAACAATGTGTGCACTGTGTGCTGCGCACCGCAACAACAGCGCTCTCATG 3108
OY 2821 GTGGAATTGACATCGCAAGTGAAGATCACTTTGTGTGAAACCTGCTGAGAAATTTCAAA 2880
DB 3109 GTGGAATTGACATCGCAAGTGAAGATCACTTTGTGTGAAACCTGCTGAGAAATTTCAAA 3168

```

```

OY 2881 AACAGAACGGGTGGCAGAAAGCTGTGGTGTGTTTCAAACTTCTGCTGTCTTCTAC 2940
DB 3169 AACAGAACGGGTGGCAGAAAGCTGTGGTGTGTTTCAAACTTCTGCTGTCTTCTAC 3228
OY 2941 AAATCACACGAGACAAATCATCCCTTGTGCGAGCTGCTCTGTGTGCTTACCTGCTAC 3000
DB 3229 AAATCACACGAGACAAATCATCCCTTGTGCGAGCTGCTCTGTGTGCTTACCTGCTAC 3288
OY 3001 ATCCCTCTGAGTCCGAGAAATCCAGAAAGCTACGTGTGTAAGTGAAGTGAAGTGC 3060
DB 3289 ATCCCTCTGAGTCCGAGAAATCCAGAAAGCTACGTGTGTAAGTGAAGTGAAGTGC 3348
OY 3061 CACGTCTACTTCACTGAGCGGAGAAAGCGAGTACAGTTCGAAAGTGAAGTGAAGTGC 3120
DB 3349 CACGTCTACTTCACTGAGCGGAGAAAGCGAGTACAGTTCGAAAGTGAAGTGAAGTGC 3408
OY 3121 CGCAGTGCACACAGCTCTGCTCTGCGAGCCCAAGTGTGAGCCACAAAGAGTCTTGTG 3180
DB 3409 CGCAGTGCACACAGCTCTGCTCTGCGAGCCCAAGTGTGAGTCAACAAAGAGTCTTGTG 3468
OY 3181 TATGTATGGCCGAGACACTGCTTCCGAGTGTGCTGCTTCTGTGAAGACGTTTCTTT 3240
DB 3469 TATGTATGGCCGAGACACTGCTTCCGAGTGTGCTGCTTCTGTGAAGACGTTTCTTT 3528
OY 3241 CTTCTGTATTAATGAAGCTGTGTAATAATTAACACCTGTCTGAATAATCAAAATCATG 3300
DB 3529 CTTCTGTATTAATGAAGCTGTGTAATAATTAACACCTGTCTGAATAATCAAAATCATG 3588
OY 3301 CCCAGAGCTCTCTGTCTTCAACAGCGCGGTTTTTAAACCCGACCTCTCAAGGTTTGA 3360
DB 3589 CCCAGAGCTCTCTGTCTTCAACAGCGCGGTTTTTAAACCCGACCTCTCAAGGTTTGA 3648
OY 3361 TGAACAGCGCTCCACCTTCACTTCTGAGATCCGCTGGGGGGGCTTTTAAAGCTAG 3420
DB 3649 TGAACAGCGCTCCACCTTCACTTCTGAGATCCGCTGGGGGGGCTTTTAAAGCTAG 3708
OY 3421 CCAGTATTAACATTTGCAAT 3442
DB 3709 CCAGTATTAACATTTGCAAT 3730

RESULT 4
ID AAA08582
AC AAA08582; standard; DNA; 4687 BP.
XX
XX
XX AAA08582;
AC
XX
XX 19-JUL-2000 (first entry)
DT
XX
XX
DE Human cytoskeleton associated protein 2 (CYSKP-2) coding sequence.
XX
XX Cytoskeleton associated protein; CYSKP-2; cancer; proliferative;
XX autoimmunity; inflammatory; vesicle trafficking; neurological;
XX cardiovascular; cell motility; reproductive; muscle disorder; ss.
XX
XX Homo sapiens.
OS
XX
XX
XX Key Location/Qualifiers
XX CDS 90..3227
XX FT /tag= a
XX FT /product= "CYSKP-2"
XX
XX PN MO200017355-A2.
XX
XX PD 30-MAR-2000.
XX
XX PF 17-SEP-1999; 99WO-US021565.
XX
XX PR 18-SEP-1998; 98US-0172226P.
XX PR 27-APR-1999; 99US-0131321P.
XX
XX PA (INCYTE) INCYTE PHARM INC.
XX

```

PI Lal P, Tang YT, Yue H, Hillman JL, Bandman O, Cortley NC;
PI Guegler KU, Patterson C, Azimzai Y, Baughn MK;
XX
DR WPI: 2000-283582/24.
DR P-PSDB; AAY91947.

Human cytoskeleton associated proteins, used to treat cell proliferative,
PT autoimmune/inflammatory, vesicle trafficking, neurological, cell
PT motility, reproductive and muscle disorders.

Claim 9, Page 101-102; 113p; English.

XX
XX
XX AAA0581-96 encode human cytoskeleton associated proteins 1 to 16 (CYSKP-
CC 1 to CYSKP-16) respectively. The sequences can be used to treat and
CC diagnose cancer and cell proliferative, autoimmune/inflammatory, vesicle
CC trafficking, neurological, cardiovascular, cell motility, reproductive
CC and muscle disorders. Pharmaceutical compositions containing CYSKP-2 to
CC CYSKP-16 can be used to treat or prevent disorders associated with
CC decreased expression or activity of CYSKP (claimed), for example,
CC atherosclerosis, cirrhosis, hepatitis, myelofibrosis, psoriasis, cancers,
CC acquired immunodeficiency syndrome (AIDS), Crohn's disease, diabetes,
CC mellitus, Goodpasture's syndrome, osteoarthritis, scleroderma and trauma.
CC CYSKP antagonists can be used to treat or prevent a disorder associated
CC with increased expression or activity of CYSKP (claimed)

XX
XX
SQ Sequence 4687 BP; 1179 A; 1261 C; 1196 G; 1051 T; 0 U; 0 Other;

Query Match 99.9%; Score 3437.2; DB 3; Length 4687;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3439; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCGCGAGCGCGCGCGCTGTGAGATATCTCTAAGCCGCTTTCATCATGGAGAAATA 60
Db 42 CCCCCAGCGCGCGCGCTGTGAGATATCTCTAAGCCGCTTTCATCATGGAGAAATA 101
QY 61 GAGCAGAGCCGAGCCGAGATCAGACTGAGGAGCCCGGAAATTCGGGATCAGTACC 120
Db 102 GAGCAGAGCCGAGCCGAGATCAGACTGAGGAGCCCGGAAATTCGGGATCAGTACC 161
QY 121 TTGGAACTGTGACAGAAAGCCGCCCAACACTTCAGGAAACTGTGTCTCAATAATC 180
Db 162 TTGGAACTGTGACAGAAAGCCGCCCAACACTTCAGGAAACTGTGTCTCAATAATC 221
QY 181 CAGATGCTGATGACACCCGAGAGGATTTGAAGTTCCAGAAAGCTCTGGAAGGTG 240
Db 222 CAGATGCTGATGACACCCGAGAGGATTTGAAGTTCCAGAAAGCTCTGGAAGGTG 281
QY 241 CTGCTGATGACATTTGCAACCACTCACTGTGGAAGTGAATTTTGGCTCGAG 300
Db 282 CTGCTGATGACATTTGCAACCACTCACTGTGGAAGTGAATTTTGGCTCGAG 341
QY 301 TTTCTGATCAAAAAAGTACAGGTGTGGCTGATCTCTTAAACCCTTTGTGAACG 360
Db 342 TTTCTGATCAAAAAAGTACAGGTGTGGCTGATCTCTTAAACCCTTTGTGAACG 401
QY 361 ATTGAAGGCCAAAGACGTTGTTAAGTTGTGTAATTTCTTCCGCTGACAC 420
Db 402 ATTGAAGGCCAAAGACGTTGTTAAGTTGTGTAATTTCTTCCGCTGACAC 461
QY 421 ACACAATCTCAAGAACTCAAGATACCTGTTGCGCTGCAAGTGAAGGACGATTTG 480
Db 462 ACACAATCTCAAGAACTCAAGATACCTGTTGCGCTGCAAGTGAAGGACGATTTG 521
QY 481 GTTCAAGGAGGTGACGTGTATGACACACAGCGACGCTCTTGAATTTCAACATTTG 540
Db 522 GTTCAAGGAGGTGACGTGTATGACACACAGCGACGCTCTTGAATTTCAACATTTG 581
QY 541 CAATCTGAGATTGGGATTTTATGTAAGCTTGGACAGAGACACTTAGCAAAAATATA 600
Db 582 CAATCTGAGATTGGGATTTTATGTAAGCTTGGACAGAGACACTTAGCAAAAATATA 641
QY 601 TACATACCTCAGACGACACTAGAGCAAAATCTGTGAATTTCAACATTAACCAATT 660

Db 642 TACATACCTCAGACGACACTAGAGCAAAATCTGTGAATTTCAACATTAACCAATT 701
QY 661 GACAAACACACGACGATTCAGATTTCCAGCTCTTGAAGATTTCCGCTGATAGATG 720
Db 702 GACAAACACACGACGATTCAGATTTCCAGCTCTTGAAGATTTCCGCTGATAGATG 761
QY 721 TATGAATCCGTTGACCCGCGCAAGGACAGGAAAGGACAGAAATCAATCTGGCCGT 780
Db 762 TATGAATCCGTTGACCCGCGCAAGGACAGGAAAGGACAGAAATCAATCTGGCCGT 821
QY 781 GCCAACACGGAATTCAGTGTTCAGGATTTCACTAAGATCAATGCTTCACTGGGCC 840
Db 822 GCCAACACGGAATTCAGTGTTCAGGATTTCACTAAGATCAATGCTTCACTGGGCC 881
QY 841 AAGTGGGAAAGCTGAGCTTCAAGGAAAGGCTTTTCATCAACACTCCGCAATGCG 900
Db 882 AAGTGGGAAAGCTGAGCTTCAAGGAAAGGCTTTTCATCAACACTCCGCAATGCG 941
QY 901 AATAGTGTACAGATACCTTGAATTCCTGATGCGCAGTCGGATTTTGCAGATGCC 960
Db 942 AATAGTGTACAGATACCTTGAATTCCTGATGCGCAGTCGGATTTTGCAGATGCC 1001
QY 961 TTTGGAATATCTGTGTTGAACATCATGCTTTTGAACCTTTTGAAGGCCAATCCA 1020
Db 1002 TTTGGAATATCTGTGTTGAACATCATGCTTTTGAACCTTTTGAAGGCCAATCCA 1061
QY 1021 AAGCCCAAGCCGCTCTTTAAGCCGGGGTCATCATTTGGTTAGTGTGCACTAG 1080
Db 1062 AAGCCCAAGCCGCTCTTTAAGCCGGGGTCATCATTTGGTTAGTGTGCACTAG 1121
QY 1081 AAGCAGTTCTGACATATGTTAAGAAAGAGACATAAGAAAGTGCAGTTTGAAGAAG 1140
Db 1122 AAGCAGTTCTGACATATGTTAAGAAAGAGACATAAGAAAGTGCAGTTTGAAGAAG 1181
QY 1141 CACAGCAATTCATTTATTCGAGGCTTGCTTCAACGCTTCAAGAACTGAATTCGAA 1200
Db 1182 CACAGCAATTCATTTATTCGAGGCTTGCTTCAACGCTTCAAGAACTGAATTCGAA 1241
QY 1201 GTGCTGAGAGCTCTCAGACAGACACAGCTTATTTGAAGAGGTGCCAATCTCCA 1260
Db 1242 GTGCTGAGAGCTCTCAGACAGACACAGCTTATTTGAAGAGGTGCCAATCTCCA 1301
QY 1261 GGGGGCCAGAGCTGCGCGCGAGAAAGGAACTTTCGCGCGGGAGCGCGGAGTGC 1320
Db 1302 GGGGGCCAGAGCTGCGCGCGAGAAAGGAACTTTCGCGCGGGAGCGCGGAGTGC 1361
QY 1321 CACCGAGCCCTGCGCGCGAGAAAGGAACTTTCGCGCGGGAGCGCGGAGTGC 1380
Db 1362 CACCGAGCCCTGCGCGCGAGAAAGGAACTTTCGCGCGGGAGCGCGGAGTGC 1421
QY 1381 TCGGGGCCACGAGAGAGAGAGAGGTCGTTAAGATTAAGACCCAGACAGTAACT 1440
Db 1422 TCGGGGCCACGAGAGAGAGAGAGGTCGTTAAGATTAAGACCCAGACAGTAACT 1481
QY 1441 CAGCCCCGCGACCAAGACAGAGCTCCCTGACTGCACTTTCGAGAGTGTCT 1500
Db 1482 CAGCCCCGCGACCAAGACAGAGCTCCCTGACTGCACTTTCGAGAGTGTCT 1541
QY 1501 GTGAATCTGCAAGGGGAGTGGCCCTTGCACAAGTGAATTTTCCAACTTGAGCCCC 1560
Db 1542 GTGAATCTGCAAGGGGAGTGGCCCTTGCACAAGTGAATTTTCCAACTTGAGCCCC 1601
QY 1561 GACACCAAGAGGAGCTCTCCCTGATGACGCGCGTGAATGAACAGGCGTGGCCCGG 1620
Db 1602 GACACCAAGAGGAGCTCTCCCTGATGACGCGCGTGAATGAACAGGCGTGGCCCGG 1661
QY 1621 ACGACGATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATTAAGCGTACTTCATA 1680
Db 1662 ACGACGATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATTAAGCGTACTTCATA 1721
QY 1681 GCTAAGGAAGTGTCTTACACGAGCGAACTATCTGAAGATCTGAAAGTTATCACTTGG 1740

Db 1722 GCTAAGAAAGTGTCTACCAACGAGCGAACATATCTGAGAGATCTCGAAGTTATCACTTCG 1781
Qy 1741 TGGTTTCAGAGCACAAGAGCAAAAGGAGCGCATGCGGAGCACTGAAAGTTCATTA 1800
Db 1782 TGGTTTCAGAGCACAAGAGCAAAAGGAGCGCATGCGGAGCACTGAAAGTTCATTA 1841
Qy 1801 TTCCCGAATTTTGAACCTTTGCACAAAATTTCACTAATTTTCTCAAGAAATTTGAGCA 1860
Db 1842 TTCCCGAATTTTGAACCTTTGCACAAAATTTCACTAATTTTCTCAAGAAATTTGAGCA 1901
Qy 1861 CGACTTGGCTGTGGGAGAGCGCTCAATATGCCAAATCAAGATTAACAAAGATGCGC 1920
Db 1902 CGACTTGGCTGTGGGAGAGCGCTCAATATGCCAAATCAAGATTAACAAAGATGCGC 1961
Qy 1921 GAGTATAGTGAAGAACTTACGGGATGAGACACTGCGGCTCACTCTGGAAGCAC 1980
Db 1962 GAGTATAGTGAAGAACTTACGGGATGAGACACTGCGGCTCACTCTGGAAGCAC 2021
Qy 1981 AGCGAGGCTTGGAGGCGCTGAGAAATGGAATCAAGACTCCCGGCGCTGAGAACTTC 2040
Db 2022 AGCGAGGCTTGGAGGCGCTGAGAAATGGAATCAAGACTCCCGGCGCTGAGAACTTC 2081
Qy 2041 TGCAGAGACTTTGAGCTGCAAGAAAGTGTGTACTTACCTTCAACCTTCTCTGCGG 2100
Db 2082 TGCAGAGACTTTGAGCTGCAAGAAAGTGTGTACTTACCTTCAACCTTCTCTGCGG 2141
Qy 2101 CCATTCACACCGGCTCATGCACTAACAGAGTCTTGAAGGCGTGTGCAAAACACCCG 2160
Db 2142 CCATTCACACCGGCTCATGCACTAACAGAGTCTTGAAGGCGTGTGCAAAACACCCG 2201
Qy 2161 CGGAGCAGCGCCGACTTCAAGGAGACTGCGAGCGCTTTGGCAGAGATCAAGAGATGAGT 2220
Db 2202 CGGAGCAGCGCCGACTTCAAGGAGACTGCGAGCGCTTTGGCAGAGATCAAGAGATGAGT 2261
Qy 2221 GCAAGCTCCACCGGTAAGATGATCAAGAGAGAAATTTCCAGAAAGTGCAGAACTCAAG 2280
Db 2262 GCAAGCTCCACCGGTAAGATGATCAAGAGAGAAATTTCCAGAAAGTGCAGAACTCAAG 2321
Qy 2281 AAAAGATTTGATTTGGCATTTGACATCTTTGTGTGTTCCGGGAAAGGAGTTTCACTCTGCGG 2340
Db 2322 AAAAGATTTGATTTGGCATTTGACATCTTTGTGTGTTCCGGGAAAGGAGTTTCACTCTGCGG 2381
Qy 2341 AGCTCAGCAAGGCTCTCGGGGAAAGGAGCTCCAGACAGGCACTGTTCTTCTGTTCAAGAC 2400
Db 2382 AGCTCAGCAAGGCTCTCGGGGAAAGGAGCTCCAGACAGGCACTGTTCTTCTGTTCAAGAC 2441
Qy 2401 GTCTTGTCTATACAGAGCGGGGGCTGACGGCTCCAAATCAATGTTTAAAGTCCACGGGAC 2460
Db 2442 GTCTTGTCTATACAGAGCGGGGGCTGACGGCTCCAAATCAATGTTTAAAGTCCACGGGAC 2501
Qy 2461 CTCTCCGCTCTATTTGGCATTTGAGAGAGAGGAAAGAGAGTGGGGGGTGCCTCCACTGC 2520
Db 2502 CTCTCCGCTCTATTTGGCATTTGAGAGAGAGGAAAGAGAGTGGGGGGTGCCTCCACTGC 2561
Qy 2521 CTGACCTCTCGGGGGCCAGGCGGAGTCCATCATGCTGGCGCGCACTTCTGCTGCGAGAT 2580
Db 2562 CTGACCTCTCGGGGGCCAGGCGGAGTCCATCATGCTGGCGCGCACTTCTGCTGCGAGAT 2621
Qy 2581 GAGAAAGTGGATTGAGAGCATTCAGATGGCCATTGACTTGGCGGAGAGAGACAGACCCG 2640
Db 2622 GAGAAAGTGGATTGAGAGCATTCAGATGGCCATTGACTTGGCGGAGAGAGACAGACCCG 2681
Qy 2641 GCGCCCTGAGTTCTTGGCCAGACAGCCCTCTGACAAAGTCCCTCTGATGAGACACCGCG 2700
Db 2682 GCGCCCTGAGTTCTTGGCCAGACAGCCCTCTGACAAAGTCCCTCTGATGAGACACCGCG 2741
Qy 2701 GCTTACACAGAGTCAAGAGATGACTGAGGCGCTCGGCGACATCGCTGAGCGCCAGGCG 2760
Db 2742 GCTTACACAGAGTCAAGAGATGACTGAGGCGCTCGGCGACATCGCTGAGCGCCAGGCG 2801
Qy 2761 CGGCAACCGGCGCAACAATGTGTGCACTGTGTGCGGACCGCAACAACAGGCTCTCATG 2820
Db 2802 CGGCAACCGGCGCAACAATGTGTGCACTGTGTGCGGACCGCAACAACAGGCTCTCATG 2861

Qy 2821 GTGACTTCAGCATCGCAGTGGAGATCAATTGTCTGAAAACCTGTGAGAAATTCAA 2880
Db 2862 GTGACTTCAGCATCGCAGTGGAGATCAATTGTCTGAAAACCTGTGAGAAATTCAA 2921
Qy 2881 AACAGCAACGGGTGGAGAAAGCTGTGGTGTGTTCAAACTTGTCTTCTTAC 2940
Db 2922 AACAGCAACGGGTGGAGAAAGCTGTGGTGTGTTCAAACTTGTCTTCTTAC 2981
Qy 2941 AATTCACACGAGAACATATCCCTTGGCAGGCTCTGTGTGCGGATCTGCTCAC 3000
Db 2982 AATTCACACGAGAACATATCCCTTGGCAGGCTCTGTGTGCGGATCTGCTCAC 3041
Qy 3001 ATCCCTCTGAGTCCGAGAAATCCAGAAAGATCAAGTGTCAAGTGCACCTCAAGTCC 3060
Db 3042 ATCCCTCTGAGTCCGAGAAATCCAGAAAGATCAAGTGTCAAGTGCACCTCAAGTCC 3101
Qy 3061 CACGTCTACTTCTTCAAGGCGGAAAGCGAGTACACGTTGAAAAGTGAATGATC 3120
Db 3102 CACGTCTACTTCTTCAAGGCGGAAAGCGAGTACACGTTGAAAAGTGAATGATC 3161
Qy 3121 CGCAGTCCACACGCTCTGCTGCGACCCACAGTGTGAGCCACAAAGTCTTGTG 3180
Db 3162 CGCAGTCCACACGCTCTGCTGCGACCCACAGTGTGAGTCAAAAGTCTTGTG 3221
Qy 3181 TATTGATGGCGGACACACTCGTTTCGGAGTGGCTGTTCTCTGGAAGACGTTCTTT 3240
Db 3222 TATTGATGGCGGACACACTCGTTTCGGAGTGGCTGTTCTCTGGAAGACGTTCTTT 3281
Qy 3241 CTCTGTATTTAATGAAGCTGTGTAATTAACACTGTGTGAATCAAAATCATGGCTT 3300
Db 3282 CTCTGTATTTAATGAAGCTGTGTAATTAACACTGTGTGAATCAAAATCATGGCTT 3341
Qy 3301 CCAGAGAGCTCTCTGTCTCTCACAGCGCGTGTTTTAAACCCGAGCTCTCAGGTTGA 3360
Db 3342 CCAGAGAGCTCTCTGTCTCTCACAGCGCGTGTTTTAAACCCGAGCTCTCAGGTTGA 3401
Qy 3361 TGAACAGCGCTCCACACTGACAGTCCGAGTCCGAGTGGGGGGCTGTTCTTATGCTAGT 3420
Db 3402 TGAACAGCGCTCCACACTGACAGTCCGAGTCCGAGTGGGGGGCTGTTCTTATGCTAGT 3461
Qy 3421 CCAGTATTTAAACATTGTCAAT 3442
Db 3462 CCAGTATTTAAACATTGTCAAT 3483

RESULT 5
AAS64760
ID AAS64760 standard; cDNA; 3187 BP.
XX
AC AAS64760;
XX
DT 13-FEB-2002 (first entry)
XX
XX
DE DNA encoding novel human diagnostic protein #564.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT.

QY 1620 GACGGAAGATGAGGAGGCGGAGGAGAGATTCCCAATGATTAAGCGTACTTCAT 1679
 DB 1621 GAGGAGAGATGAGGAGGCGGAGGAGAGATTCCCAATGATTAAGCGTACTTCAT 1680
 QY 1680 AGCTAAGAGAGTGTACCAACCGAGGAAATATCTGAAGATCTGAAATTATCACTTC 1739
 DB 1681 AGCTAAGAGAGTGTACCAACCGAGGAAATATCTGAAGATCTGAAATTATCACTTC 1740
 QY 1740 GTGGTTTCAGAGACAGTGAAGCAAGAGAGCGCAATGCGGAAAGCACTGAAAAGTCTCAT 1799
 DB 1741 GTGGTTTCAGAGACAGTGAAGCAAGAGAGCGCAATGCGGAAAGCACTGAAAAGTCTCAT 1800
 QY 1800 ATTCCGGAATTTTGAACCTTTGCAAAATTTCTAATTAATTTCTAAGGAATTTAGCA 1859
 DB 1801 ATTCCGGAATTTTGAACCTTTGCAAAATTTCTAATTAATTTCTAAGGAATTTAGCA 1860
 QY 1860 ACGACTTGGCTGTGGGAGGCGCTCAATGCGCAATCAGAGATTACCAAGAAATCGG 1919
 DB 1861 ACGACTTGGCTGTGGGAGGCGCTCAATGCGCAATCAGAGATTACCAAGAAATCGG 1920
 QY 1920 CGATGTCAATGCTGAAGAACATTCAGGAGCATGAAGCACTGGCGGCTCACTGTGAAGCA 1979
 DB 1921 CGATGTCAATGCTGAAGAACATTCAGGAGCATGAAGCACTGGCGGCTCACTGTGAAGCA 1980
 QY 1980 CAGCGAGGCTTGGAGAGGCGCTGAGAAATGGAATCAAGAGCTCCGCGGCTGAGAACTT 2039
 DB 1981 CAGCGAGGCTTGGAGAGGCGCTGAGAAATGGAATCAAGAGCTCCGCGGCTGAGAACTT 2040
 QY 2040 CTGCAAGAGATTTGAGCTGAGAGAGTGTGTTACTACCGCTCAACACTTCTCTGCG 2099
 DB 2041 CTGCAAGAGATTTGAGCTGAGAGAGTGTGTTACTACCGCTCAACACTTCTCTGCG 2100
 QY 2100 GCGACTGCAACCGGCTCATGCACTACAGAGAGTCTGAGAGCGGCTGTGCAACCAACCC 2159
 DB 2101 GCGACTGCAACCGGCTCATGCACTACAGAGAGTCTGAGAGCGGCTGTGCAACCAACCC 2160
 QY 2160 GCGGAGGCGAGCGGCTTGAAGGAGCTGCGGAGCGGCTTGGGAGAGATCAACGAGATGGT 2219
 DB 2161 GCGGAGGCGAGCGGCTTGAAGGAGCTGCGGAGCGGCTTGGGAGAGATCAACGAGATGGT 2220
 QY 2220 GGCACAGCTCCACGATGATGATCAAGATGAGAAATTTCCAGAACTGCAACCACTCAA 2279
 DB 2221 GGCACAGCTCCACGATGATGATCAAGATGAGAAATTTCCAGAACTGCAACCACTCAA 2280
 QY 2280 GAAAGATTTGATTGGCATTGACAAATCTTGTGTTCCGGGAAAGGAGTTCACTCGTGGG 2339
 DB 2281 GAAAGATTTGATTGGCATTGACAAATCTTGTGTTCCGGGAAAGGAGTTCACTCGTGGG 2340
 QY 2340 CAGCCCTCAGACAGCTCTCGGGGAAAGGCGCTCCAGAGCGGCAATGTTCTCTGTTCAACA 2399
 DB 2341 CAGCCCTCAGACAGCTCTCGGGGAAAGGCGCTCCAGAGCGGCAATGTTCTCTGTTCAACA 2400
 QY 2400 CGTCTCTATATACAGAGCGGAGGAGCTGACGCGCTCAATCAAGTTAAATCCACGGGCA 2459
 DB 2401 CGTCTCTATATACAGAGCGGAGGAGCTGACGCGCTCAATCAAGTTAAATCCACGGGCA 2460
 QY 2460 GCTCCCGCTCTATGCGATGACGATTGAGAGAGCGAAGACGAGTGGGAGGCTGCCACTG 2519
 DB 2461 GCTCCCGCTCTATGCGATGACGATTGAGAGAGCGAAGACGAGTGGGAGGCTGCCACTG 2520
 QY 2520 CCGTGAACCTCCGCGGAGCGGAGCTGACATCGTGGCGGCAAGTTCTCGGTCCAGAT 2579
 DB 2521 CCGTGAACCTCCGCGGAGCGGAGCTGACATCGTGGCGGCAAGTTCTCGGTCCAGAT 2580
 QY 2580 GGAGAAGTGGGTTGAGAGATCCAGATGGCCATTTGACTGGCGGAGAGAGAGAGAGAGAGAGAG 2639
 DB 2581 GGAGAAGTGGGTTGAGAGATCCAGATGGCCATTTGACTGGCGGAGAGAGAGAGAGAGAGAGAG 2640
 QY 2640 CCGCCCTGAGTTCTTGGCGAG 2699
 DB 2641 CCGCCCTGAGTTCTTGGCGAG 2700

QY 2700 GGCTGACAGAGATGAGAGATGAG 2759
 DB 2701 GGCTGACAGAGATGAGAGATGAG 2760
 QY 2760 CCGGACCGCGGAGACAAATGAGTGAAGTGTGCTGGGACACCGGACAGCAGCGTCTCAT 2819
 DB 2761 CCGGACCGCGGAGACAAATGAGTGAAGTGTGCTGGGACACCGGACAGCAGCGTCTCAT 2820
 QY 2820 GGTTGACTTCAAGATGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 2879
 DB 2821 GGTTGACTTCAAGATGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 2880
 QY 2880 AAACAGCAACGAGTGAAGAGAGTGTGGTGGTGTGCAAACTTGTGCTGTCTTCTA 2939
 DB 2881 AAACAGCAACGAGTGAAGAGAGTGTGGTGGTGTGCAAACTTGTGCTGTCTTCTA 2940
 QY 2940 CAAATGACACAGAGACAAATCATCCCTTGGCAGCCTGCTGTGCTGTGCTGTGCTGTGCT 2999
 DB 2941 CAAATGACACAGAGACAAATCATCCCTTGGCAGCCTGCTGTGCTGTGCTGTGCTGTGCT 3000
 QY 3000 CATCCCTCTGAGTCCGAGAAATCCAGAAAGACTACGTTGCACTGCACTTCAAGTC 3059
 DB 3001 CATCCCTCTGAGTCCGAGAAATCCAGAAAGACTACGTTGCACTGCACTTCAAGTC 3060
 QY 3060 CCAGCTCTACTCTTCAAGGCGGAGAGAGTACAGTTCAGAAAGGTGATGAGATGAT 3119
 DB 3061 CCAGCTCTACTCTTCAAGGCGGAGAGAGTACAGTTCAGAAAGGTGATGAGATGAT 3120
 QY 3120 CCGCAGTGCACACAGCTCTGCTCGGAGCCGACCGTGTGAGGACCAAGAGTCTTGT 3179
 DB 3121 CCGCAGTGCACACAGCTCTGCTCGGAGCCGACCGTGTGAGGACCAAGAGTCTTGT 3180
 QY 3180 GTATTGA 3186
 DB 3181 GTATTGA 3187

RESULT 6
 AAC98992
 ID AAC98992 standard; cDNA, 3094 BP.
 XX
 AC AAC98992;
 XX
 DT 09-MAR-2001 (first entry)
 XX
 XX Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:220.
 DE
 XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection;
 KW diagnosis; identification; cytostatic; neuroprotective; nootropic;
 KW immunomodulatory; relaxant; contraceptive; gynaecological;
 KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
 KW linkage analysis; tissue identification; tissue typing; forensic; neural;
 KW immune system; muscular; reproductive; gastrointestinal; pulmonary;
 KW cardiovascular; renal; proliferative; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO20005320-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005989.
 XX
 PR 12-MAR-1999; 99US-0124270P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-579444/54.
 DR P-PSDB; AAB54227.
 XX
 PT New nucleic acid that is a pancreatic cancer antigen for preventing,

PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition.
XX
PS Claim 1, Page 664-665; 1379pp; English.

XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC gynecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridization probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 3094 BP; 736 A; 849 C; 755 G; 749 T; 0 U; 5 Other;

Query Match 49.1%; Score 1688.6; DB 3; Length 3094;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1699; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1740 GTGGTTTCAGACACAGTGAAGAGACGCCATGCGGAAGCATTAAGTCTCAT 1799
DB 62 GTGGTTTCAGACACAGTGAAGAGACGCCATGCGGAAGCATTAAGTCTCAT 121
QY 1800 ATTCCGAAATTTGAACCTTGACAAATTTGATCTAATTTTCCAAAGAAATGAGA 1859
DB 122 ATTCGGAATTTGAACCTTGACAAATTTGATCTAATTTTCCAAAGAAATGAGA 181
QY 1860 ACGACTTGCCCTTGAGGAAGCGCGTCAAAATGCCAAATCAGAGATTACAAAGATCGG 1919
DB 182 ACGACTTGCCCTTGAGGAAGCGCGTCAAAATGCCAAATCAGAGATTACAAAGATCGG 241
QY 1920 CGATGTCATGCTGAAGAAATTCAGGGCATGAAAGCATCTGGCGCTCACTGTGGAAGCA 1979
DB 242 CGATGTCATGCTGAAGAAATTCAGGGCATGAAAGCATCTGGCGCTCACTGTGGAAGCA 301
QY 1980 CAGCGAGGCTTGAGAGCCCTGGAAGATGGAATCAAGAGCTCCCGCGGCTGGAAGACTT 2039
DB 302 CAGCGAGGCTTGAGAGCCCTGGAAGATGGAATCAAGAGCTCCCGCGGCTGGAAGACTT 361
QY 2040 CTGCAAGACCTTTGAGCTGCAAGAGGTGTGTTACCTACCGCTCAACACCTTCTCTCGG 2099
DB 362 CTGCAAGACCTTTGAGCTGCAAGAGGTGTGTTACCTACCGCTCAACACCTTCTCTCGG 421
QY 2100 GGCATCTGACCGGCTCATGCACTAACAGAGGTCTCTGAGCGGCTGTGAAACACCAACC 2159
DB 422 GGCATCTGACCGGCTCATGCACTAACAGAGGTCTCTGAGCGGCTGTGAAACACCAACC 481
QY 2160 GCGAGGCGCACGCGGCACTTCAGGGAAGTCCGAGCGCTTGAGAGATCAAGAGATGAGT 2219
DB 482 GCGAGGCGCACGCGGCACTTCAGGGAAGTCCGAGCGCTTGAGAGATCAAGAGATGAGT 541
QY 2220 GGCACAGCTCAACGATGATCAAGATGGAATTTTCAAGAGCTGCAAGAACTCAA 2279
DB 542 GGCACAGCTCAACGATGATCAAGATGGAATTTTCAAGAGCTGCAAGAACTCAA 601
QY 2280 GAAAGATTTGATTTGGCATTTGATTTGTTGTTCCGGGAAGGAGTTCACTCGTCTGGG 2339
DB 602 GAAAGATTTGATTTGGCATTTGATTTGTTGTTCCGGGAAGGAGTTCACTCGTCTGGG 661

QY 2340 CAGCTCAGCAAGCTCTGCGGGGAAGGGCTCCAGACGGGCAATGTTCTTCTGTTCAACGA 2399
DB 662 CAGCTCAGCAAGCTCTGCGGGGAAGGGCTCCAGACGGGCAATGTTCTTCTGTTCAACGA 721
QY 2400 GCTCTGCTATACACAGACCGGGGGCTGACGGCTTCAATCAATGTTAAAGTTCACGGGCA 2459
DB 722 GCTCTGCTATACACAGACCGGGGGCTGACGGCTTCAATCAATGTTAAAGTTCACGGGCA 781
QY 2460 GCTCCCGGCTATGAGCAATGGAAGAGCAACAGATGGGGGGTGGCCCACTG 2519
DB 782 GCTCCCGGCTATGAGCAATGGAAGAGCAACAGATGGGGGGTGGCCCACTG 841
QY 2520 CCTGACCTCCGGGGGCGACGGGAGTCCATGATCTGAGCCGCGCAATTCGGTCCGAGAT 2579
DB 842 CCTGACCTCCGGGGGCGACGGGAGTCCATGATCTGAGCCGCGCAATTCGGTCCGAGAT 901
QY 2580 GAGAAATGGGTTGAGACATTCAGATGCGCAATGACCTGGCGGAAGAGCAGACGCC 2639
DB 902 GAGAAATGGGTTGAGACATTCAGATGCGCAATGACCTGGCGGAAGAGCAGACGCC 961
QY 2640 CGCCCTGATTTCTGGGCGACAGCCCTTGAACAAGTCCCTGATGAAGCAGCCG 2699
DB 962 CGCCCTGATTTCTGGGCGACAGCCCTTGAACAAGTCCCTGATGAAGCAGCCG 1021
QY 2700 GAGCTGACAGAGTCAAGAGATGACCTGAGCGCTCGGCGCATGCTGAGCGCGCAGG 2759
DB 1022 GAGCTGACAGAGTCAAGAGATGACCTGAGCGCTCGGCGCATGCTGAGCGCGCAGG 1080
QY 2760 CCGGACCGCGGCAACATATGTCACGTGTCTGAGCAACGCAACACACGCTTCCAT 2819
DB 1081 CCGGACCGCGGCAACATATGTCACGTGTCTGAGCAACGCAACACACGCTTCCAT 1140
QY 2820 GGTGACCTTCAGCATCGAGTGGAGAAATCAGTTGCTGGAACCGCTGAGGAATTCAA 2879
DB 1141 GGTGACCTTCAGCATCGAGTGGAGAAATCAGTTGCTGGAACCGCTGAGGAATTCAA 1200
QY 2880 AAACAGCAACGGGTGGGAGAGCTGTGGGTGTTCACAACTTCTGCTGTTCTTA 2939
DB 1201 AAACAGCAACGGGTGGGAGAGCTGTGGGTGTTCACAACTTCTGCTGTTCTTA 1260
QY 2940 CAATCAACAGAGCAATCAATCCCTTTCAGAGCTGCTGCTGCTACTCGCTCAC 2999
DB 1261 CAATCAACAGAGCAATCAATCCCTTTCAGAGCTGCTGCTGCTACTCGCTCAC 1320
QY 3000 CATCCCTCTGAGTCCGGAATCATCAGAAAGATCAAGTTTCAAGCTGCACTTGAAGTC 3059
DB 1321 CATCCCTCTGAGTCCGGAATCATCAGAAAGATCAAGTTTCAAGCTGCACTTGAAGTC 1380
QY 3060 CCACGCTACTACTTCAGGGGCGAAAGCGAGTACAGTTTCGAAAGGTGAGTGAAGTAT 3119
DB 1381 CCACGCTACTACTTCAGGGGCGAAAGCGAGTACAGTTTCGAAAGGTGAGTGAAGTAT 1440
QY 3120 CCGCAGTGCACACAGCTTGTCTGCGGACCCCAAGTGTGAGGCCAAGAGTCTCTTGT 3179
DB 1441 CCGCAGTGCACACAGCTTGTCTGCGGACCCCAAGTGTGAGGCCAAGAGTCTCTTGT 1500
QY 3180 GATATGATGGGCGGACACACTGTTTCGAGTGGCTCTTTCGGAAGAGTTCCTT 3239
DB 1501 GATATGATGGGCGGACACACTGTTTCGAGTGGCTCTTTCGGAAGAGTTCCTT 1560
QY 3240 TCTTCTGATTAATGAGCCTGTGTAATTAACCTGTCTGAAATCAAAAACATGGCT 3299
DB 1561 TCTTCTGATTAATGAGCCTGTGTAATTAACCTGTCTGAAATCAAAAACATGGCT 1620
QY 3300 TCCACAGAGCTCTCTGTCTCTCAACAGCGGCTTTTAAACCCCGACCTCTCAGCGTTGA 3359
DB 1621 TCCACAGAGCTCTCTGTCTCTCAACAGCGGCTTTTAAACCCCGACCTCTCAGCGTTGA 1680
QY 3360 ATGAACAGGCTCCACCTCAAGTCTGAGATCCGCTGGGGGGGCTGTTCTTTAGCTAGT 3419
DB 1681 ATGAACAGGCTCCACCTCAAGTCTGAGATCCGCTGGGGGGGCTGTTCTTTAGCTAGT 1740

QY 3420 GCCAGTATTAAACATTGCATT 3442
|||||
Db 1741 GCCAGTATTAAACATTGCATT 1763
|||||

RESULT 7
ADQ84807
ADQ84807 standard; cDNA; 3997 BP.

AC ADQ84807;
XX
DT 07-OCT-2004 (first entry)
XX

Human tumour-associated antigenic target (TAT) cDNA sequence #1621.

human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
cancer; cell proliferative disorder; gene; ss.

Homo sapiens.

MO2004060270-A2.

22-JUL-2004.

15-OCT-2003; 2003WO-US029126.

18-OCT-2002; 2002US-0418988P.

(GETH) GENENTECH INC.

(WUTD/) WU T D.

(ZHOU/) ZHOU Y.

Wu TD, Zhou Y;

WPI; 2004-534300/51.

New nucleic acid molecule and encoded polypeptide, for diagnosing,
preventing or treating cell proliferative disorders such as cancer.

Claim 1; SEQ ID NO 1621; 5504bp; English.

The present invention describes an isolated tumour-associated antigenic
target (TAT) nucleic acid comprising: (a) any of 4632 nucleotide
sequences (see SEQ ID NO.1 to 4632); (b) the full-length coding region of
(a); (c) the complement of (a) or (b); (d) a sequence that has 80%
sequence identity to (a)-(c); or (e) a sequence that hybridises to (a) -
(c). Also described: (1) an expression vector comprising the above
nucleic acid; (2) a host cell comprising the above expression vector; (3)
a process for producing a polypeptide; (4) an isolated polypeptide
comprising: (a) an amino acid sequence encoded by any of the above
nucleotide sequences; (b) an amino acid sequence encoded by the full-
length coding region of the above nucleotide sequences; or (c) a sequence
having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
comprising the above polypeptide fused to a heterologous polypeptide; (6)
an isolated antibody that binds to the above polypeptide; (7) a process
for producing the antibody; (8) an isolated oligopeptide that binds to
the above polypeptide; (9) a tumour-associated antigenic target (TAT)
binding organic molecule that binds to the above polypeptide; (10) a
composition of matter comprising the above (chimeric) polypeptide,
antibody, oligopeptide or TAT binding organic molecule, in combination
with a carrier; (11) an article of manufacture comprising a container and
the composition of matter contained within the container; (12) methods of
inhibiting the growth of a cell that expresses the above protein, where
the growth of the cell is at least in part dependent upon a growth
potentiating effect of the above protein; (13) a method of
therapeutically treating a mammal having a cancerous tumour comprising
cells that express the above protein; (14) a method of determining the
presence of a protein in a sample suspected of containing the protein
described above; (15) methods of diagnosing the presence of a tumour in a
mammal; (16) a method for treating or preventing a cell proliferative
disorder associated with increased expression or activity of the above
protein; and (17) a method of binding an antibody, oligopeptide or
organic molecule to a cell that expresses the protein described above.

CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.

XX
SQ Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;

Query Match 30.1%; Score 1037.4; DB 12; Length 3997;
Best Local Similarity 60.2%; Pred. No. 1e-270;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

QY 69 GCCGACCCGAGATCAACGACTGGGGCCCCGGAATAATGGGGATAGTACTGGAAGC 128
|||
Db 150 GCAGACTGCAAGGATGCGCTGGGGTCCAGACCCCTGGGGATTAGACCTTTAGCC 209
|||
QY 129 TGGACAGAAAGCCGCCCAACCTTCAGGAAATCTGTCCATCAAAATCCAGATGCT 188
|||
Db 210 TGGGACAGACTCTCTTCCGAGAAATGCAAGAAAGACCTGCACCTCAGATTAAGCTGCT 269
|||
QY 189 GATGACACCCAGAGAGCAATTTGAAGTCCCAAGAGCTCTGGGAAGTGTCTGGA 248
|||
Db 270 GCACAAACACCATGGAATATTGACATTGAGCTTAATGCGATGCGCAGATTACTGAC 329
|||
QY 249 TGCAGTTTGCACACCTCAACTCGTGAAGGTGACTATTTTGGCTCGAGTTCTGA 308
|||
Db 330 ACAAGTGTGAAGCGTTTAACTGTGTAAGTGACTACTTGGAGTGAAGTTTCAAAA 389
|||
QY 309 TCACAAAAGATCAACGATGCTGGCTGATCTCTTAAACCAATTTGTGAACAGATTGAAG 368
|||
Db 390 TACTAGTCCCTATGATATTGGCTTGAACCTATGAACCATCATATTAGCAATAGGAAG 449
|||
QY 369 GCCAAGACAGTGTGTTGTTAAGTTGTGTAATTTCTTCCGCTGACCAACACAACT 428
|||
Db 450 GCCAAGAAATGTGTGCTTCCCTAGCTGTAATAATTTTCCACATCTGCTGAGCT 509
|||
QY 429 CCAAGAAAGATCAACAGTACTGTTGCGCTGACAGTGAAGAGAGATTTGCTCAAG 488
|||
Db 510 ACAAGAAAGATTAACAAGTACTTGTTCCTTGAACCTTAAGAGAGACTGTGGAAG 569
|||
QY 489 CAGTTGAGGTGTAATGATGACACGACGACTCTTGTGATTTTCAACATTTGTAATCTGA 548
|||
Db 570 GCGTTTGACTGTGTGATGACACGACGCGCTTCTCAGTCCCATCTCTGAGTCCGA 629
|||
QY 549 GATTGGGATTTTGTATGAAGCTTGGACAGAGACCTTAGCAAAAATAATATACATACC 608
|||
Db 630 AATAGAGATTAAGATGAAGACCTGGAACGAGGACCTCAAGTGAACGAGTATTTGCC 689
|||
QY 609 TCAGCAAGACGACATGAGAGACAAATCGTGAATTTTCACTAACCATTTGAGCAAAAC 668
|||
Db 690 TGGCCAGACGACTGCTTGAAGAAATTAAGATTTCAAGAAAGACGATGGGCCAGAC 749
|||
QY 669 ACCAGAGATGATGATTTTCCAGCTCTTAGAGATTTGCGCTGAGATGATGAT 728
|||
Db 750 ACTGTGTGATGATTTTCCAGCTCTTAGAGATTTGCGCTGAGATGATGAT 809
|||
QY 729 CCGGTTGACCCGGCCAAAGAGAGAGAGACAAAGATCAATCTGCGCTTGCACAAAC 788
|||
Db 810 CAGATTTCACTAGCTTCTGACAGGGAAGAAACAAAGATTTCACTGCGGAGTTTCCACAT 869
|||
QY 789 GGAATTTCAAGTTTTCAGAGTTTCACTAAGATTTTCACTGCGGAGTTTCCACAT 848
|||
Db 870 GGGTGTACTGTGTTTCAAGGGAACCAACAAATTTCACTGCGGAGTTTCCACAT 929
|||
QY 849 GAGTGAAGCTTCAAG 908
|||
Db 930 TAACTGATCTTCAAG 989
|||
QY 909 GTACCAAGATTTGATTTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 968
|||
Db 990 TTACAGGACATTAAGATTTTGTGGTGTAGAGATGATGATGATGATGATGATGATGAT 1049
|||

QY	969	AATCTGTGTTGAACATCATG	CGCTCTCTTTAGACCTTTTGAAGAGCCCAACCAAGCCCA	1028
Dp	1050	GATTTGTGTGAGATACAC	CTTTTATTTTGAAGCTTTGACCAACCTTAAGCCAAAGCAAA	1109
QY	1029	GCCGCTCTCTTTAGCCGGGGGATCATTTT	CGGTTGAGTGTGGAGCTGAGAGCAGT	1088
Dp	1110	AGCGGCTCTTTCACGCCGGGGCTCTCTCTT	CAGATACAGTGAAGAACTCAGAAACA	1169
QY	1089	TCTGACTATGTTAAAGAAAGAGCA	TTAAGAAAGGTGCAAGTTTGAAGAAAGACAGCA	1148
Dp	1170	AGTAGATTATTTCAAAAGCA	GAGTGAATGAAGAAATTCATATGAAGAAAGCACAGCA	1229
QY	1149	GATTTCATTCTATCCGGAGC	CTTGCTTACAGCTTACAGAACTGAATTTGGAAAGTCTGGA	1208
Dp	1230	GACCCACACGTCC-----	GTTTCGAGCTTGTGACTGACAGCTTACCA	1271
QY	1209	GCAGTCTCAGAGAGCA	CAGCCTTCACTTTTGAAGAAAGGTGCGAATTCAGAGGGGCA	1268
Dp	1272	ACAGAGCATCTCATTTCCCGAGGGATTTGAGGAC	CTCTGCTTCCTCCCATTTTCAAGGAATGC	1331
QY	1269	GAGCTGCGCGAGAAAGAAAGAA	CGAAGGTTTCCGCGCGGGAGCGGGAGTGCACCCGAG	1328
Dp	1332	CTTTTATCTCGCTCTCTCTCTCTCCCTC	CACTCTGTCTCCCTCTGCGCTGCAAGATTTAAGACAG	1391
QY	1329	CCCTGCGCGAGAGAAAGCCCGCGGGTAA	CAAGCAGGCGGAGCCGCTCTGGCGCC	1388
Dp	1392	CAGCAGCTCCCTCAGATCCCCAGAGTTT	CCTAAGTCAAGAGTCCAGCTGCAGAGAGCGC	1451
QY	1389	CACGAGAGAAAGAGAGAG---TCGTTAA	AGATAGACCCAGCAGATAA	1445
Dp	1452	CAGTGAAGCAATGTGTGAGGCGCCCGACAC	CAATCGGCCAGGCCCTCGGGCCCCCGC	1511
QY	1446	CCCCGAGCAAGCAACAGGCTCCTGTGAC	TGGCAGTCTGTCTGTGA	1505
Dp	1512	ACTCGAGCTGTGTCCAGGCTTTTCCAGAA	AGTCTCAGGCTTCTCCCTCAGACCGGAA	1571
QY	1506	CTGCGAGGGGGAAGTGCGCCCTGCCA	AGTCACTTGTCTCCCACTGAGCCCCGAC	1565
Dp	1572	GAGCCCCCTGAGTCTGAGCCCTCGCAT	TTTCAAGTGCCTT-----TGCGCCAGCTGA	1622
QY	1566	CAAGCAGGCTCTCTCCCTTGA	TTACAGCGCGCTGTGAATGACCAAGCCTGCCCGGAGCGGA	1625
Dp	1623	ACAGGGCTCATCTCCCATCTCTGAGCC	CTGTCTTCAAGTATCTGAGCGAGCGGGAATGA	1682
QY	1626	CGATGAGGATGAGGGCGGAGGAAGAT	TCCCACTGAATTAAGCTTCTCATAGCTAA	1685
Dp	1683	CTGCGAG---GAGCCCA	CAACAGGCGGTCTGCAAGCAGAGCTTCTCATAGTCA	1739
QY	1686	GGAAGTGTCTTACCA	CCGAGCGAATATCTGAAGATCTCGAAGTTATCACTTGTGTGT	1745
Dp	1740	AGAGATTCTCGCTACAGAACGA	ATACCTCAAGGATTTAAGATTTATTCAGTGTGTT	1799
QY	1746	TCAGAGCACTGAGCAAAAGAGAG	AGCCATCTCCGGAAGCACTGAAGAACTCTCATATTTCC	1805
Dp	1800	CCGAGCGCAATGTGTGAAGAGAG	CGCCATCTGCGACTGTGATGACGTCTCTTCTC	1859
QY	1806	GAATTTTGAACCTTTGCA	CAATTTTCAATTAATTTTCTCAAGGAAATGAGACAGACT	1865
Dp	1860	CAATCATGATCCCATCTTATG	ATGCTTCAAGAGGCTTCTGTGCGAGGTGAGCAGAGCT	1919
QY	1866	TGCCCTGTGGAAAGCGCGCTCAAA	TGCCCAATCAGAG---ATTACCAAAAGATCGGCGA	1922
Dp	1920	GGCACTCTGGGAAGGGGCTTCAAA	AGCCCAACAAAGGCAGTATCAAGCAATCGGGGA	1979
QY	1923	TGTCACTGTGAAGAA	CATTCAAGGGCATGAGCACTGTGGCGCTCACTGTGGAAGCACAG	1982
Dp	1980	CATCTGTCTCAGGAACATG	CGCCAGTTTAAAGAGTTTACAGCTTCTTCCAAACATGA	2039
QY	1983	CGAGGCTCTTGAAGCGCTTGA	GAATGGAATCAAGAGCTCCGGGGGCTCGAGAACTTCTGTG	2042
Dp	2040	CGAGGTCTTACAGAACTGGA	AAAGGCTTCAACAAAGCTGTGAAGATTTGAGGCGCAGTGTGA	2099
QY	2043	CAGAGACTTTGAGCTGCA	GAAGGTGTGTATCTTACCGCTTCAACACTTCTCTGCGGCT	2102

Db	2100	CAAGAGTTTGAAGTCAGAAAGGCTCTCTACTTTCCTCTCAACAGTTTCTTCTGAAGCC	2159
Qy	2103	ACTGCAACGGGTCTATGCACTACAAAGAGAGTCTTGAGACGGCTGTGCAAAACACACCCGC	2162
Db	2160	CATCCAGGGGTGTGTGCACTACCGGCTGTGTGTGGCGGCTTATGGGGACATTACAGCC	2219
Qy	2163	GAGCCACGCCGACTTTCAGGGAGCTGCGGAGCCGCTTTGGCAGAGATCACGGAGATGTGGC	2222
Db	2220	CGGGCACCATTGACTACCTGTGACTGCGCTCATGACGCGCTTAAAGCCATACAGAGGTGACAC	2279
Qy	2223	ACAGCTCCACGGTTCAGATGATCAAGAGTGGAGAAATTTCAGAGGTGACCGAATCAAGAA	2282
Db	2280	CACACTACAGCACTTCTCATTCGCGCTGGAGAACTTCAGAGAGTTAACGAGCTTCAGCG	2339
Qy	2283	AGATTGATTTGGCAATTGACAACTTTGTGTGTCCGGGAAGGAGTTCAATCGCTGTGGCAG	2342
Db	2340	GGAGCTGTGGGCACTAAGAACTCATTTGCTCTCCGGAGAGGAGTTCAATCCGTAGGGCTG	2399
Qy	2343	CCTGAGCAAGCTCTCGGGGAGGGGCTCCAGACGCGATGTTCTTCTGTTCACAGAGT	2402
Db	2400	CCTTCACAAAGTCCACAAAGGGCGCTGACAGAGAGATGTTTTTCTGTCTCAGATAT	2459
Qy	2403	CTGTCTATACACGAGCCGGGGGCTGACGGCTCCAACTCAATTAAATCCACGGGCAAGT	2462
Db	2460	GTTTCTGTACCAAGCAAAAGAGTGTGAGGAGCCAGCACTTCCGAAATCCGGGGCTTCCT	2519
Qy	2463	CCGCGTCTATGGCATYGCATTTGAGAGAGCGAAGCAGATGGGGGGTGTGCCCCCTGCT	2532
Db	2530	TCCCTTCCAAAGGCACTGTCTGTGAAAGAAATGATATACGATGTGTCTTTTCCACACTTTT	2579
Qy	2533	GACCCCTCGGGGCGCAGCGGAGTCCATCATCTGTGGCGCGAGTTTCTCGTTCGAGATGA	2582
Db	2580	CACCATCTACGCGGCTCAGAAAACAACTGATGTGGCAGCAGCACTCGGCTGGAGAAAGA	2639
Qy	2583	GAATGTGGTTAGGACATCCAGATGGCCATTGACTGTGGCGAGAAAGACAGAGCCCCGC	2642
Db	2640	GAATGTGATGTGGAACCTGAATCTCCGGGATCCAAAGCCAAAGATGGGCGGTGACAGGC	2699
Qy	2643	CCCTGAGTTCCTGGCCAGCA-----GCCGCCCTGACAAACAAAGTCCGCTGATGAGCCAC	2696
Db	2700	CCCTGCACTGCGAGCCGCACTGTGTGTGCACTGTCTCCCCCAGATCCCCCAACAGATATC	2759
Qy	2697	CGCGGCTGACCAAGAGTCAAGAGATGAACCTGAGCGCTCGCGCACATCTGCTGAGCGCA	2756
Db	2760	TCTG---GAGCAGGAGTCAAGAAATGAATGTCTGGGGGTGTCCGACGTTCCCTGAGGGGCA	2816
Qy	2757	GGCCCGGACCGCGGCAACACATGATGTGCACTGTGTGGCACCGCAACACAGCTCTC	2816
Db	2817	TGGCAGACCGGGGCCAAACCAACAAAGCACTGTGTGTGTAACGGAAACACAGCGTGT	2876
Qy	2817	CATGTGTGAATTTTCAGATTCGCAATGCAAGTCAAGTTCTCTGGAAACCTGTGAGGAATTT	2876
Db	2877	CAGGGCAGACCAACGTGCAAGTGTGTGAAACCACTTTCAAGATATTTGCTTAGAAAGTT	2936
Qy	2877	CAAAAACGACCAACGATGAGTGTGTGAAAGGCTGTGGGGTGTCTTTACAACTTGTGTTCTT	2996
Db	2937	CTACAAATCACACAGACAAATCATCCCCCTTGGCAGCTGTGCTCTGCTCGGCTAATCGCT	2996
Qy	2937	CTACAAATCATCATAGATGACTACCCATCGGCGAGGCTCCCGCTGTGGGCTTCAACGCT	3056
Db	2997	CACATATCCCTCTGAGTCCGAGAAATCCCAAGAAAGATTAGTGTCAAGCTGCACCTTCAA	3056
Qy	2997	GAGCATCCCAAGGAGGCGCATGCAATACAAAGACTTAATGTTTTCAAGCTCCAGTTCAA	3116
Db	3057	GTCCCAAGTCTACTACTTCAAGGCGGAGAAACGATACAGTTTCGAAAGGTGATGAGAT	3116
Qy	3117	ATCCCAAGTCTACTTCTTCGGGGCTGAGACAGATGACAAATTTGAAAGGTGAGAGGT	3176
Db	3117	GATCCGAGTGCACAGCTCTGCTTCGACAGCCC	3151

1446 CCCGAGCCAAAGCAGAGCTCCCTGACTGAGTCTCTCACTTTCGAGAGCTGTGTGAA 1505
1512 ACTCCAGCTTGATCCAGGCTTTCCAGAAAGATCTTACGCTTCTCCCTCCAGCCGGA 1571
1506 CTGCGAGGGGGAGTGGCCCTGGCCAACTGTACCTTGTCTCCCAACTGAGCCCGAC 1565
1572 GAGCCCTGAGTGTGAGCCCTGCAATTCAGGTGCTT-----TGGGCCCAAGCTGA 1622
1566 CAGCAGAGCTCTCTCTTATATGAGCCGCTGTGAAATGACAGAGCTGCCCCCGACGA 1625
1623 ACAGGGCTATCCCACTCTGAGCCCTGTCTCAATGATGCTGGGAGAGCGGAGTGA 1682
1626 CGATGAGATGAGGGCCGAGGAGAAAGATTTCCCACTGAATAAGCCTACTTATAGCTAA 1685
1683 CTGCGAG---GAGCCAGACACAGAGGCGTGTCTGAGACGAGGCTTACTTATAGTCA 1739
1686 GGAAGTGTCTACCAAGCAGGAAATATCTGAAGATCTGAAAGTTATCACTTCTGTGTT 1745
1740 AAGAGTTCCTGCTACAGAAAGAACTCAAGATTCAGAGTATTAAGTTATACCGTGTGTT 1799
1746 TCAGAGCAGAGTGAAGAAAGAGAGCCATGCGGAGCACTGAAAGTCTCATATTTCC 1805
1800 CCGCAGCCGAGTGTGAAGAGAGAGCGCATGCTGCACTGTGATGACGCTGCTCTTC 1859
1806 GAATTTGAACTTTGACAAATTTCTATCTAATTTCTCAAGAAATTTGACCAAGACT 1865
1860 CAACATCATCCCATCTATGATTCACAGAGGCTTCTGCGGAGGTGAGACAGAGGCT 1919
1866 TGCCTGTGGGAAGCCGCTCAATGCCCAATCAGAG--ATTACAAAGATGGCGA 1922
1920 GGCACCTGTGGGAAGGCGCTCCAAAGCCACAAAGGCACTCATCAAGATCGGGGA 1979
1923 TGTCACTGAGAAATTCATCAAGGCAATGAAGCACTGGGCGTCACTGTGAGACAG 1982
1980 CATCTGCTCAGAAATGCGCCAGTTAAAGAGTTTACAGCTTCTCAAGAAATGA 2039
1983 CGAGGCTTGAAGGCGCTGGAATGGAATCAAGAGCTCCGCGGCTGGAACCTTCTG 2042
2040 CGAGGCTTAACAGAACTGGAAGGCTACCAAGCTGTAAAGAGTTGAGAGCAAGTGA 2099
2043 CAGAGACTTTGAGCTGAGAGAGGTGTACTTACCTGCTCAACACTTCTCTGCGGCT 2102
2100 CAAGGAGTTGAGCTGCAAGAGGTGTGCTTCAACACTTCTCTGCTGAGGCT 2159
2103 ACTGCAACGAGCTCATGCACTAACAAGGCTCTGAGGCGGTGTGAAACACACCGGCT 2162
2160 CATCCAGCGGCTGTGCACTACCGCTGTGCTGCGCGCCCTATGCGGACATTTACAGCCC 2219
2163 GAGCCAGCGGCACTTCAGGGAATGCGAGCGCTTTGAGAGAGATCAGAGATGCTGCG 2222
2220 CGGGGACCATGACTAGCTGAGCTGCCATGACGCGCTGAAGGCATACAGAGGTGACAC 2279
2223 ACAAGCTCCAGGTAAGATGATCAAGATGAAGATTTCCAGAAAGCTGACAGAACTAAGAA 2282
2280 CACACTACAGCAATTCATCTATCGGCTGGAAGAACTGCAAGAGCTAACGAGACTGAGG 2339
2283 AGATTTGATTGGGATTTGACATCTTGTGTTCCGGGAAGGAGTTATCCGTCTGGGAG 2342
2340 GGAAGCTGTGGGATTAAGAACTTCTCTCTGCGAGGAGTTTATCCGTGAGGAGCTG 2399
2343 CCTCAGCAAGCTCTCGGGGAGAGGAGCTCCAGCAGCGCATGTTCTTCTGTTCACGACGT 2402
2400 CTTTCAAGAGCTCAACAAAGAGGCTGACAGAGAGAGATGTTTTTTCTGTTCAGATAT 2459
2403 CCTGTATACAGAGCCGGGGGCTGACGGCTTCATCTCAATGCTTAAAGTCCAGGGGAGCT 2462
2460 GTTGTGTGTACAGAGAAAGAGTGTGACAGGACCACTTCGGAGATCGGGGCTCTCT 2519
2463 CCGGCTCTATGGATGAGATTGAGAGAGAGCAAGTGGGGGAGTCCCACTGCGCT 2522
2520 TCCCTCTCAAGGAGTGTGTGTGAGAAAGATTAAGAGTGTCTGTTCACACTGTTT 2579

2523 GACCTCCGGGGCCAGCGGAGTCCATCATGCTGCGCCGCAAGTTCTGATCGAGATGGA 2582
2580 CACATCTAACGGGCTCAGAAAACAAATGTTGTGGCAGCCAGCACTCGGCTGAGAAAGA 2639
2583 GAAATGGTTGAGGACATCCAGATGGCCATTGACTGTGCGGAGAAAGAGCAGAGCCCGC 2642
2640 GAAATGGATGTGAGACTGAACTCCGATCCCAAGCAGCAGAAAGATGGCGGTGACACGCG 2699
2643 CCTGAGTTCTGGCAGCA-----GCCCCCTGACAAACAAGTCCCTGATGAAGCCAC 2696
2700 CCTGCACTGCGAGGCGGACCTGTGTGACATGTGTCCCCCAATCTCCCAAGAGATATC 2759
2697 CCGGCTGACCAAGAGTGAAGATGATGACGCTGTGCGCATGCTGCTGAGCGCCA 2756
2760 TCTG---GAGCAGAGTGAAGATGATGCTGGGGGTGTCCGCAAGCTCCCTGAGAGGGCA 2816
2757 GGGCCCGCAGCCGGCCAAACAATGATGATGTGTGCTGACACCGGCAACACAGGCTCTC 2816
2817 TGGCAGACACCGGGCCAAACCAATGACGTTGTGTGATCTGGAAACACAGCGTGTG 2876
2817 CATGCTGACTTCAGCATGCGAGTGAAGATCAATGTTGTCTGGAACCTGTGAGAAATT 2876
2877 CAGGCAAGACCAAGTGAAGTGTGAGAACCAAGCTTTCAGGATATCTGTAAAGAAATT 2936
2877 CAAAAACAGCAACGGGTGGCAGAACTGTGGTGTTCACAACTTCTGCTGTCTT 2936
2937 CAAAAACAGTATGGCTGGCAGAAAGCTCTGGGTGTGTACCACTTCTGTTTGTCTT 2996
2937 CTACAAATCAACACAGACATCATCCCTTGTCCAGGCTGTGCTGTGCTATCTGCT 2996
2997 CTACAAACTCATTCAGATGATGACTACCACTGGCCAGCTCTCCGCTGTGGCTTACAGCT 3056
2997 CACATCCCTCTGTGAGTCCGAGAACATCAGAAAGACTAGTGTCAAGCTGCACTTCA 3056
3057 GAGCATTCGCCAGGAGGCGGATGGATACAAAGACATATGTTTCAAGCTTCAAGTTCA 3116
3117 ATCCAGCTCTTACTTCTCCGGCTGAGAGCAAGTACATTTGAAAGTGTGATGAGGT 3176
3117 GATCCGAGTGCACCAAGCTTGTGCTCGGACCC 3151
3177 GATCCAGGGGCGCAGCAGCTCAGCCGAGGAGGCC 3211

RESULT 9
AD564452
ID AD564452 standard; DNA; 3997 BP.
XX
AC AD564452;
XX
AC
XX
DT 16-DEC-2004 (first entry)
XX
DE Human KIAA0793 gene.
XX
XX Human; interferon; IFN; gene; ds; KIAA0793.
XX
OS Homo sapiens.
XX
PN US2004185489-A1.
XX
XX
PD 23-SEP-2004.
XX
XX
PF 17-MAR-2004; 2004US-00802432.
XX
XX
PR 21-MAR-2003; 2003EP-00006263.
XX
XX
PA (CERT/) CERTA U.
XX
PA (FOSE/) FOSE S.
XX
PA (WEYE/) WEYER K.
XX
PI CertA U, Foser S, Weyer K;
XX

DR MPI; 2004-689187/67.
XX GENBANK; AB018336.
PT Determination of biological activity of compound which can modulate gene
transcription, involves contacting host with compound, determining
PT general transcriptional gene response of host, and quantitating gene
response induced by compound.
XX
PS Claim 8; SEQ ID NO 16; 53bp; English.
CC The invention relates to a method for determining the biological activity
of a compound which can modulate gene transcription. The method involves
CC contacting a host with a compound, determining the general
transcriptional gene response of the host and quantitating the gene
response induced by the compound. This gene transcription assay method is
useful for the determination of the biological activity of a compound
CC which can modulate gene transcription. The present sequence is the human
KIAP0793 gene. This sequence is used to estimate the influence of the
CC peptidation site on the transcriptional activity of interferon (IFN) by
measuring this IFN-induced gene expression pattern of the melanoma cell
line ME35 via the oligonucleotide array technology.
XX
SQ Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;
Query Match 30.1%; Score 1037.4; DB 13; Length 3997;
Best Local Similarity 60.2%; Pred. No. 1e-270;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;
69 GCGGACCCGAGATCGACTGGGGGCCCCGAAAAATTCGGGAGTCACTTGGAAACG 128
150 GCGACCTGCGAGGATCGCTTGGGTGCCGACCCCTGTGGAGTTAGCACTTGGAGCC 209
129 TGGACGAAGCCGCCCAACACCTTCAAGAAATCTGTCTCATTAATTCAGATGCT 188
210 TGGGCGACTCTCTTCCCGAGATGCAAGAGACCTGCACCTGAGTTAAAGCTGCT 269
189 GGATGACCCGAGGAGCATTTGAAGTTCCAAAGAGCTCTGGAGAGGTGCTGTGGA 248
270 GGAACACCATGGAATATTTTACATTGAGCTTAATGGATGGCAGGATTAATCTGAC 329
249 TGCAGTTTCACACCACTCACTCGTGAAGGTGACTATTTTGGCTCGAGTTTCTGA 308
330 ACAGGTGGAAGCGTTTAACTGTGAGATGTGACTACTCGGAGTGAAGTTTCAAAA 389
309 TCACAAAAAGATACGCTGTGCTGATCTCTAAACCATTGTGAACAGATTGGAAG 368
390 TACTGCTCTATGATTTGGCTTGAACCTATGAACCCATCATTAAGGAATAGGAAG 449
369 GCCAAGCAAGTGTGTTAAAGTTTGGTGAATCTTTCGCTGACACACACACT 428
450 GCCAAGAAATGTGTGCTTGCCTAGCTTAAATTTTTCACCTGATCTGTGTAGCT 509
429 CCAGAGAACTCACAAGGTAAGTTCGCTGAGTGAAGAGAGAGAGAGAGAGAGAG 488
510 ACAGAGAAATATACAAAGTACTTGTTCGCTGCAACTTAAGAGAGAGAGAGAGAG 569
489 CAGGTGAGCTGTATGACACAGCGAGTCTTGAATTTTCAACATTTGTGAATCTGA 548
570 GCGTTTGAAGCTGTGAGACCAAGCGGCTTTCACAGTCCCACTCTGAGAGTGA 629
549 GATGGGAGTTTGTGAAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 608
630 AATGAGAGATTAAGATGAACGCTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 689
609 TCAGCAAGACGACTAG 668
690 TGGCCGAGCACTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 749
669 ACCAGAGAGATGAGATTTTCAAGCTCTTGAAGATTTGCGCTGAGAGAGAGAGAT 728
750 ACTGTGTAGTGGATTTTCCAGGTGCTCGAAATTTGCTGAAGATTTGAATGAGCAG 809
729 CCGGTTGACCCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 788

810 CAGATTTCAATGCTTTCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 869
789 GGGATTTCTAGGTTTTCAGGGTTTCACTAAGATCAATGCTTCACTGCGCCAGAGTGC 848
870 GGGTGTACTGTGTTTTCAGGGGACCAACAAATTCACATTTTCACTGCTCCAGGCTCG 929
849 GAAGCTGAGCTTCAAG 908
930 TAACTAAGCTTCAAG 989
909 GTACAGAGATTCCTGAAATTCCTGATGCGAGTCCGAGATTTCTGAGATCTTTCGAA 968
990 TTAACAGAGACATTAAGATTTTGTGGTGTGATGATGATGATGATGATGATGATGAT 1049
969 AATCTGTGTTGAACATGAGCTTCTTGAATCTTTCAGATTTTGAAGAGAGAGAGAGAG 1028
1050 GATTTGTGTGAGTATCAACACTTTTGTAGACTTTTGAAGAGAGAGAGAGAGAGAGAG 1109
1029 GCGCGCTCTTGAAGCGGGGGTCAATTTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1088
1110 AGCGTCTTTCAGCGGGGCTCTCTTCAAGATCAAGTGAAGAGAGAGAGAGAGAGAGAG 1169
1089 TCTGACTATGTTAAAG 1148
1170 AGTAAATTTTTCAG 1229
1149 GATTCATTTATTCGAGAGCTTGTTCACAGCTTACAGAGAGAGAGAGAGAGAGAGAG 1208
1230 GACCCACAGTCC-----GTTCAAGCTTGAAGTGCAGAGAGAGAGAGAGAGAGAG 1271
1209 GGAAGTTCAG 1268
1272 ACAGAGCATCTAATTCCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1331
1269 GAGCTGCGCGAG 1328
1332 CTTTATCTGCT 1391
1329 CCCTGCGCGAG 1388
1392 CAGCAGCTCTCTACAGATCTCCAGAGGTTTCTAAGTCAAGATCTCAGCTGAGAGAGAG 1451
1389 CACGAG 1445
1452 CAGTGAAG 1511
1446 CCCGAGCAAG 1505
1512 ACTCAGAGCTGTGATCAGAGCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1571
1506 CTCGAG 1565
1572 GAGCCCTGAGTGTGAG 1622
1566 CAGCAGAGCT 1625
1623 ACAGAGCTATCCCACTCTAGAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1682
1626 CGATGAGATGAG 1685
1683 CTGCGAG---GAGCCAG 1739
1686 GGAAGTGTCTACACCGAG 1745
1740 AAGAGATTTCTGCTTACAG 1799
1746 TCAGAGCAG 1805
1800 CCGCAGCGAG 1859
1806 GAATTTGAAGCTTTGCAAAATTTTCAATTTTCTCAAGAGAGAGAGAGAGAGAGAGAGAG 1865

Db 1860 CAACATCCATCCCATCTATGATTCACAGAGGCTTCTGCGCAGAGTGAACAGAGCT 1919
Qy 1866 TGGCCCTGGGGAAGGCGGCTCAAAATGCGCAATCAAG--ATTACCAAGAAATGGGGA 1922
Db 1920 GGACACTGGGAAGGGGCTCCAAAGGCCACAAAGGAGCTCATCAACGATGGGGA 1979
Qy 1923 TGTCACTGGAAGAACTTCAGGGGATGAAGCACTGCGGCTCACTGTGAGACAG 1982
Db 1980 CATCTGCTCAGAAACATGCGCAAGTTAAAGAGTTTACAGCTACTTCCAAAGACATGA 2039
Qy 1983 CGAGGCTTGAAGGCCCTGGAATGAATCAAGAGCTCCGCGGCTGGAAGACTTCTG 2042
Db 2040 CAGAGCTCTAAGAACTGGAAGAGCTCAAAAGCTGTAAAGAGTGAAGCAGAGTGA 2099
Qy 2043 CAGAGACTTGAAGTGAAGAGTGTATACCAACGCTCAACCTTCTCTGCGGCT 2102
Db 2100 CAAAGGTTTGAAGTGAAGAGTGTCTGCTTCTCAACAGTTCTCTCTGAAGCT 2159
Qy 2103 ACTGCAACGGCTCATGACATCAAGCAGGTCCTGAGCGGCTGTGCAACACACCGGCT 2162
Db 2160 CATCCAGCGGCTGCTGCTCACTACCGCTGCTGCGCGGCTATGAGCATTTACAGGCC 2219
Qy 2163 GAGCAACGCGGCTTCAAGGACTGCGGAGCCGCTTTGCGAAGATCAAGAGATGAGTGC 2222
Db 2220 CGGGGACCATGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTGACTG 2279
Qy 2223 AAGCTCCAGCGGATGATGATCAAGATGAAGATTTCCAGAAAGCTGCAAGAACTCAAGA 2282
Db 2280 CACACTACAGCAATTTCTATCCGCTGGAAGAACTGCAAGAGCTTACGAGCTGAGCG 2339
Qy 2283 AGATTGATTGGCATTTGACATCTTGTGTTCCGGAAGGAGTTTATCCGCTGAGGCGAG 2342
Db 2340 GAGCCTGCTGAGCATTAAGAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2399
Qy 2343 CCTCAGCAAGCTCTGCGGGAAGGGGCTCAAGCAGGCACTGTTCTTCTTCAAGAGCT 2402
Db 2400 CTTTCAAGAGCTCACCAAGAGGGGCTGCAAGAGAGATGTTTTTCTTCTTCAAGATAT 2459
Qy 2403 CCGTGTATCAGAGCGGGGCTGACGGCTCCATCATGTTTAAAGTCCACGCGGAGCT 2462
Db 2460 GTTGTCTTACAGAAAGAGTGTGAGGAGCAGGCCACTTCCGAGTCCGGGCGCTCT 2519
Qy 2463 CCGGCTTATGAGCATGACATTTGAGGAGAGCAGAAAGAGTGGGGGTGCCCCGCT 2522
Db 2520 TCCCTCCAAAGGCTGCTGTGGAAGAAAGTGAACGATGCTGTTTCCACAGTCTT 2579
Qy 2523 GACCTCCGCGGCGCAGCGGCACTCATCTGTCGCGGCACTTCTCGGTCCGAGATGA 2582
Db 2580 CACCATCTACCGGCTCAGAAACCAATCTGTGTGCGAGCAGACACTCGCTGAGAAAGA 2639
Qy 2583 GAAAGTGGGTGAAGCATCAGATGGCCATTGACCTGCGGAGAAAGAGCAGAGCGCCG 2642
Db 2640 GAAAGTGGATGCTGAGACTGAATCTCCGCAATCCAGAGGCGCAAGAGTGGCGGTGA 2699
Qy 2643 CCCTGAGTTCCTGCGCAGCA-----GCCCCCTGACCAAGAGTCCCTGATGAAGCCAC 2696
Db 2700 CCCTGACAGCGAGCGGCGCAGCTGTGTCAGTGTCCCCCAAGTCCCCCAAGAGATATC 2759
Qy 2697 CGCGGCTGACCGAGATCAGAGATGACCTGAGCGGCTCGCGCAGCATCGCTGAGCGGCA 2756
Db 2760 TCTG---GAGCAGAGATCAGAAATGATCTCGGGGTGTCGCCAGTCTCCCTGAGAGGGA 2816
Qy 2757 GCGCCCGCAGCGGCGGCAACAATGATGTCAGTGTGTCGCAACCGGCAACAGCGGCTGC 2816
Db 2817 TGGCCAGACCGGGCCAAACCAATGACGATGCTGTCTGTACCGGAACCAAGCGTGTG 2876
Qy 2817 CATGTGACATTCAAGCATGCAAGTGAAGATCAAGTTGTGGAACCTGCTGAGAAAT 2876
Db 2877 CAGGGGAGACCAAGTGAAGCTGTGCAAGAACAGCTTTCAAGATATCTCTAAGAAAGTT 2936
Qy 2877 CAAAAACAGCAACGGGTGCGAAGAGTGTGGGTGTGTTTCAAAACTTCTGCTGTCTT 2936
Db 2937 CAAAAACAGTCAATGGCTGGCAGAAAGCTGTGGGTGTGTTTACCAACTTCTGTTTCTT 2996

Qy 2937 CTACAAATCACACGAGCAATCATCCCTTGGCAGGCTGCTGCTGCTGCTGCTGCTGCT 2996
Db 2997 CTACAAATCATCAGATGATGATCACTCCACTGGCCAGGCTCCGCTGCTGCTGCTGCT 3056
Qy 2997 CACCATCCCTCTGAGTCCGGAACATCCAGAAAGATCAAGTGTTCAGAGCTGCACTTCA 3056
Db 3057 GAGCATCCCGCAGGAGGCGGATGGCATACAAAGATATGTTTCAAGCTCAGTTCA 3116
Qy 3057 GTCCCACTTATCTTCAAGGCGGGAAGGAGTACAGTTCGAAGAGTGAATGAAGT 3116
Db 3117 ATCCCACTTATCTTCTGCGGCTGAGAGCAAGTCAATTTGAAGAGTGAATGAAGT 3176
Qy 3117 GATCCGAGTGCACACAGCTCTGCTGCGGAGCCCC 3151
Db 3177 GATCCAGGGGCGCAGCTACAGCGGAGGGGCC 3211

RESULT 10
ACLS5653
ID ACLS5653 standard; cDNA; 637 BP.
XX
XX
ACLS5653;
XX
XX
DT 24-MAR-2005 (first entry)
XX
XX
DE Human colon cancer differentially expressed polynucleotide, SEQ ID:1788.
XX
KW Differential expression; diagnosis; therapy; drug screening; cancer;
KW neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
KW ss.
XX
XX Homo sapiens.
XX
XX MO2005000087-A2.
XX
XX PD 06-JAN-2005.
XX
PF 13-MAY-2004; 2004MO-US015421.
XX
XX PR 03-JUN-2003; 2003US-0475872P.
XX
XX PA (CHIR) CHIRON CORP.
XX
DR Randazzo F, Moler E, Escobedo J, Garcia PD;
XX
XX New isolated polynucleotides, which are differentially expressed in colon
XX cancer cell, useful for treating cancer, e.g. colon cancer, breast
XX cancer, or pancreatic cancer.
XX
XX Claim 1; SEQ ID NO 1788; 97bp; English.
XX
XX The invention relates to 9672 polynucleotides (ACLS5866-ACLS6357) which
XX are differentially expressed in colon cancer cells. The invention also
XX relates to vectors and host cells comprising a differentially expressed
XX polynucleotide of the invention; a method for detecting a cancerous cell
XX by detection of a gene product of the polynucleotides; a method for
XX inhibiting a cancerous phenotype of a cell by inhibiting a gene product
XX of the polynucleotides; a method of treating an individual with cancer by
XX administration of a modulator of a gene product of the polynucleotides;
XX and an isolated antibody that specifically binds to a polypeptide encoded
XX by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
XX antibodies, and methods are useful for the detection of cancerous cells;
XX for the diagnosis, prognosis and management of cancer; for the
XX identification of agents that modulate the phenotype of cancerous cells;
XX for the identification of therapeutic targets for cancer chemotherapy;
XX and for the treatment of cancer, especially colon cancer and metastasized
XX colon cancer, but also breast or pancreatic cancer. The polynucleotides
XX are also useful as a source of probes or primers for use in diagnostic
XX methods. The differentially expressed polynucleotides or their encoded
XX proteins can additionally be used as vaccines to modulate primary immune

CC responses for the prevention or treatment of cancer. The present sequence
CC represents a specifically claimed polynucleotide which is differentially
CC expressed in colon cancer. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

XX Sequence 637 BP; 144 A; 192 C; 153 G; 148 T; 0 U; 0 Other;

Query Match 18.0%; Score 620.6; DB 14; Length 637;
Best Local Similarity 99.4%; Pred. No. 9.1e-158;
Matches 623; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 2779 ATGTCGACGTGTGTCGTCGACGCGCAACACGAGCTCTCCATGCTGACCTTACATCGCA 2838
DB 11 ATGTCGACGTGTGTCGTCGACGCGCAACACGAGCTCTCCATGCTGACCTTACATCGCA 70
QY 2839 GTGAGAAATCAGTGTCTGGAAAACCTGCTGAGGAAATTCAAAAACAGCAACGGGTGGCAG 2898
DB 71 GTGAGAAATCAGTGTCTGGAAAACCTGCTGAGGAAATTCAAAAACAGCAACGGGTGGCAG 130
QY 2899 AAGCTGTGGGTGTGTTACAAAATTCTGCTGTTCTTCTTCAAAATCACACGAGCAAT 2958
DB 131 AAGCTGTGGGTGTGTTACAAAATTCTGCTGTTCTTCTTCAAAATCACACGAGCAAT 190
QY 2959 CATTCCCTTCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3018
DB 191 CATTCCCTTCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 250
QY 3019 AACATTCAGAAAGACTACGCTGTTCAAGCTGCACTTCAAGTCCACGCTCTACTACTTACAG 3078
DB 251 AACATTCAGAAAGACTACGCTGTTCAAGCTGCACTTCAAGTCCACGCTCTACTACTTACAG 310
QY 3079 GGGGAAAGGAGTACACGTTCCAAAGTGAATGAAAGTATCCGACGTCACACGCTCT 3138
DB 311 GGGGAAAGGAGTACACGTTCCAAAGTGAATGAAAGTATCCGACGTCACACGCTCT 370
QY 3139 GCCTGCGACCCGACGCTGTTAGCCAAAGACTCTGTGATGATGATGATGATGATGATGATG 3198
DB 371 GCCTGCGACCCGACGCTGTTAGCCAAAGACTCTGTGATGATGATGATGATGATGATGATG 430
QY 3199 CTCGTTTCCGACGCTGCTCTTCTGGAAGACGTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3258
DB 431 CTCGTTTCCGACGCTGCTCTTCTGGAAGACGTTCTTCTTCTTCTTCTTCTTCTTCTTCT 490
QY 3259 CTGTTAAATTAACACCTGTCTGAAAATCAAAAACATGCTTCCGACGCTCTCTCTGTC 3318
DB 491 CTGTTAAATTAACACCTGTCTGAAAATCAAAAACATGCTTCCGACGCTCTCTCTGTC 550
QY 3319 TCACAGACCGCGCTTTTAAACCCGACCTCTGAGCGCTTGAATGAAACAGCGCTCCGACT 3378
DB 551 TCACAGACCGCGCTTTTAAACCCGACCTCTGAGCGCTTGAATGAAACAGCGCTCCGACT 610
QY 3379 CCAAGTCTGTCATCCGCTGGGGGCGCT 3405
DB 611 CCAAGTCTGTCATCCGCTGGGGGCGCT 637
```

RESULT 11
ACD92278/C
ID ACD92278 standard; cDNA; 582 BP.

XX ACD92278;

XX 23-SEP-2003 (first entry)

DE Human colon cancer cell expressed cDNA #690.

XX Open reading frame detection; genome sequencing; colon cancer;
KM breast cancer; population genome analysis; genetic shift; cancer;
KM antibiotic resistance; antibiotic non-tolerance; congenital disease;
KM agriculture; food crop genome; resistance gene; retrovirus;
KM influenza virus; eukaryotic pathogen detection; crypanosome; Plasmodium;

KM gene; ss.

XX Homo sapiens.

XX US2002155438-A1.

XX 24-OCT-2002.

XX 27-SEP-1999; 99US-00406117.

XX 20-NOV-1998; 98US-00196716.

XX (SIMP/) SIMPSON A J G.

XX (NETO/) NETO E D.

XX (BREN/) BRENTANI R R.

XX Simpson AJG, Neto ED, Brentani RR;

XX WPI; 2003-182626/18.

PT Determining open reading frames of genome of an organism e.g. a human
PT suffering from cancer involves use of single oligonucleotide primer at
PT low stringency for preparing single-stranded cDNA from mRNA of
PT individual.

XX Example 9; Page 105; 959pp; English.

XX The invention describes a method of determining open reading frames in
XX the genome of organism, comprising contacting mRNA from cell of organism
XX with a single oligonucleotide primer (I) at low stringency, preparing
XX single-stranded cDNA by reverse transcribing mRNA with (II), amplifying
XX cDNA, sequencing the product, and repeating the contacting, preparing
XX and amplifying steps with different primers and sequencing resulting
XX nucleic acids. The method is useful for: determining that a known
XX nucleotide sequence from a genome of an organism corresponds to a
XX nucleotide sequence of an open reading frame; for preparing a contig,
XX nucleic acid molecule from a genome of an organism; and for sequencing
XX all or part of a genome of an organism. mRNA is obtained from mammalian
XX or human cell which is associated with a pathological condition e.g. a
XX colon cancer or breast cancer cell. The method is useful for analyses of
XX populations of subjects and can be used to carry out genetic analyses of
XX large or small populations. further, it can be used to study living
XX systems to determine if e.g. there have been genetic shifts which render
XX an individual or population more or less likely to be afflicted with
XX diseases such as cancer, to determine antibiotic resistance or non-
XX tolerance, and so forth. The method can also be used in the study of
XX congenital diseases, and the risk of affliction to a foetus, as well as
XX the study of whether the conditions are likely to be passed to offspring
XX through ova or sperm. The analyses for pathological conditions can be
XX carried out in all animals, plants, birds, fish, etc. Using this method,
XX in the area of agriculture, for example the genomes of food crops can be
XX studied to determine if resistance genes are present, defects in plant
XX genomes can also be studied in this way. Similarly, the method permits
XX determination of the pathogens which integrate into the genome, such as
XX retroviruses and other integrating viruses such as influenza virus, have
XX undergone shifts or mutations, which may require different approaches to
XX therapy. This method is also applied to eukaryotic pathogens, such as
XX trypanosomes, different types of Plasmodium, etc. The method essentially
XX eliminates sequencing of non-coding portions. This sequence represents a
XX polynucleotide isolated from human colon cancer cell cDNA library

XX Sequence 582 BP; 109 A; 147 C; 185 G; 140 T; 0 U; 1 Other;

Query Match 16.1%; Score 554.2; DB 10; Length 582;
Best Local Similarity 99.3%; Pred. No. 1e-139;
Matches 556; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2631 CAGCAGCCCGCGCTGAGTCTTGCCAGACGACCCCTGACACAGTCCCTGAGGA 2690

DB 571 CAGCAGCCCGCGCTGAGTCTTGCCAGACGACCCCTGACACAGTCCCTGAGGA 512

QY 2691 AGCCACCGCGGCTGACGAGAGTCAAGAGATGACTGAGCGGCTGCGCACATCGCTGA 2750

PR 20-SEP-2000; 2000US-02345052P.
PR 22-SEP-2000; 2000US-0234509Z.
PR 22-SEP-2000; 2000US-0234567B.
PR 25-SEP-2000; 2000US-0234923A.
PR 25-SEP-2000; 2000US-0234924B.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280B.
PR 26-SEP-2000; 2000US-0235637B.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863B.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0241867P.
PR 01-NOV-2000; 2000US-0245084P.

PA (AVAIL-) AVALON PHARM.
XX
XX
PI Young PE, Augustus M., Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppe DR, Weaver Z;
DR WPI; 2002-188264/24.
XX
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 6249; 44bp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in (I)
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 847 sequences (given in ABL51664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytosolic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms'
CC tumour.
XX
XQ Sequence 506 BP; 118 A; 112 C; 135 G; 141 T; 0 U; 0 Other;

Query Match	13.8%	Score 474.4;	DB 6;	Length 506;
Best Local Similarity	99.6%;	Pred. No. 4.7e-118;		
Matches 486; Conservative	0;	Mismatches	1;	Gaps 1;

QY	2824	GACTTTCAGCATCCGAGTGGAGGAATCACTGTCTTGGAAAACCTGCTGAGAGAAATTTCAAAAC	2883
Db	506	GACTTCAGCATCCGAGTGGAGGAATCACTGTCTTGGAAAACCTGCTGAGAGAAATTTCAAAAC	447
QY	2884	AGCAACGGGGTGGCGAAGCTGTGGGTGTGTTCAAAACTTCTGCTCTTCTACAA	2943
Db	446	AGCAACGGGGTGGCGAAGCTGTGGGTGTGTTCAAAACTTCTGCTCTTCTACAA	387
QY	2944	TCAACACCAAGACAAATCATCCCTTTGCCAGCTGCTCTGCTCGGCTACTGGCTACACATC	3003
Db	386	TCAACACCAAGACAAATCATCCCTTTGCCAGCTGCTCTGCTCGGCTACTGGCTACACATC	327
QY	3004	CCCTCTGAGTCCGAGAACATCCAGAAAGACTACGTGTTCAAGCTGCACCTCAAGTCCAC	3063
Db	326	CCCTCTGAGTCCGAGAACATCCAGAAAGACTACGTGTTCAAGCTGCACCTCAAGTCCAC	267
QY	3064	GTCCTACTACTTCAGGGCGGAAAGCGAGTACAGTTTGGAAAGGTGATGGAATGATCCGC	3123
Db	266	GTCCTACTACTTCAGGGCGGAAAGCGAGTACAGTTTGGAAAGGTGATGGAATGATCCGC	207
QY	3124	AGTGCACAACAGCTGTGCTCGCGACCCCAAGTGTGAGCCACAAAGAGTCTCTGTGTAT	3183
Db	206	AGTGCACAACAGCTGTGCTCGCGACCCCAAGTGTGAGTCACAAAGAGTCTCTGTGTAT	147
QY	3184	TGATGGCCGGACACACTCTGTTCCGCAAGTGTGCTTCTCGAAGACGTTTCTCTTCTT	3243
Db	146	TGATGGCCGGACACACTCTGTTCCGCAAGTGTGCTTCTCGAAGACGTTTCTCTTCTT	87
QY	3244	CTGTATTATATGAACCTGTGTAAATTTAAACAACGTGTGAAAATCAAAAACATGGCTTCCC	3303
Db	86	CTGTATTATATGAACCTGTGTAAATTTAAACAACGTGTGAAAATCAAAAACATGG-TTCCC	28
QY	3304	AGCAGCTC 3311	
Db	27	AGCAGCTC 20	

```

RESULT 15
ADQ63863
ID ADQ63863 standard; cDNA, 3200 BP.
XX
XX ADQ63863;
XX
XX 07-OCT-2004 (first entry)
XX
XX Novel human cDNA sequence #1024.
XX
XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
XX cytosolic; gene therapy; diagnostic marker; morbid state; osteoporosis;
XX neurological disease; Alzheimer's disease; Parkinson's disease; dementia
XX cancer.
XX
XX OS Homo sapiens.
XX
XX EP1440981-A2.
XX
XX PD 28-JUL-2004.
XX
XX PF 21-JAN-2004; 2004EP-00001196.
XX
XX PR 21-JAN-2003; 2003JP-00102206.
XX PR 09-MAY-2003; 2003JP-00131392.
XX
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX PI Isegaki T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Iehi S;
XX PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX WPI; 2004-535376/52.
XX DR P-PSDB; ADQ66051.
XX
XX PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,

```

PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX
 PS Claim 1; SEQ ID NO 1024; 2449bp; English.

XX The invention relates to 2495 novel polynucleotides (I) and their encoded
 CC polypeptides, sequences hybridizing to these nucleotides, sequences
 CC encoding partial polypeptides and sequences having 70% or 90% identity to
 CC the nucleotide and protein sequences. The nucleotides and polypeptides
 CC are useful as diagnostic markers or therapeutic target for the diseases
 CC or morbid states. They are also useful for treating osteoporosis,
 CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
 CC dementia and various cancers. This sequence corresponds to a nucleotide
 CC sequence of the invention.

XX Sequence 3200 BP; 951 A; 659 C; 718 G; 872 T; 0 U; 0 Other;

Query Match 13.0%; Score 447; DB 12; Length 3200;

Best Local Similarity 65.5%; Pred. No. 3.7e-110;

Matches 654; Conservative 0; Mismatches 345; Indels 0; Gaps 0;

QY 144 CCCAACAACCTTCAGAAAACCTGTCATCAAAATCCAGATCTGATGACCCAGGA 203
 DB 160 CCCAAGGCGCTGCTGAGATGCTACATTTAAAGTCAGATTTTGGATGATTCAGAA 219
 QY 204 GGCATTGAAGTCCACAAAGAGCTCCTGGAGAGTGCTGATGACAGTTTGCAACCA 263
 DB 220 GATTTTGTGTTGATCAAAAGTCATCCGGAGAGCAATTTTAACTGATGATTCAGCCA 279
 QY 264 CCTCAACCTGCTGAGAGTGATATTTTGGCTCGAGTTTCTGATCAGAAAAGATCAC 323
 DB 280 TCTAAATCTTGCGAAGAAAGATATTTTGGATTAGATTCTGCAGCATTCTGAAATAA 339
 QY 324 GGTGTGCTGATCTCTTAACCAATGTTGAACAGATTAGAGGCCAAGCAGCTTGT 383
 DB 340 TGTTTGGCTGAGCTTTTGAGCCATACAAAGCAGTAAATAATCTTAGAGAGATTGT 399
 QY 384 TGTTAAGTTTGTGTAATTTCTTCCGCTGACCAACACCACTCAAGAAAGACTCAC 443
 DB 400 TTTCAAAATTTATGTAATTTTCCAGTGAACCTCGACATCTGGGAGAAACTTAC 459
 QY 444 AAGTACCTGTTGCGCTGAGTGAGAGCACTTGGCTCAAGGAGGTTGACGTGTAA 503
 DB 460 AAGGTATCTTTTACTCTTCAATAAAGAGATTGGCTTAGGAAGGCTTCATGACAG 519
 QY 504 TGACACCAAGCGAGCTCTCTGATTTTCAACATTTGCAATCTGAGATTGGGATTTTGA 563
 DB 520 TGACACCTGTACGCTGTATGATCTCAATCTTACATCAGAACTTGGAGACTTTCA 579
 QY 564 TGAAGCTTGGACAGAGACCTTAGCAAAAATAATACATCACTCAGCAGAGCGACT 623
 DB 580 TGAAGAAACAGATAGAGAGATCTGGCAAACTCGTACTTTCAAACCAAGACTGTTT 639
 QY 624 AAGAGCAAAATGCTGAAATTTTCAACATACCAATTGGAACAAACACAGCAGATCAG 683
 DB 640 AAGAGGCAAGATCATCACTTTCATCAGAAAGCATTGGCAGAGGCCAGCTGAATCTGA 699
 QY 684 TTTCCAGCTCTTGAAGATTGCGCTGAGATGATGATGATCCGTTGCAACCCGCG 743
 DB 700 CATTCGTACTAGCATAGCAAGAGCTGATATGATGATGATGATGATGATGATGATG 759
 QY 744 CAAGGACAGGAAAGGACAGCAATCATCTGCGCTTGGCAACAGGGAATTTAGTGT 803
 DB 760 CAGTATGATGAGAGGATGAGATTCACCTGCGCTGTTGCTCAGATGAGAGTACTGATTT 819
 QY 804 TCAGGCTTCACTAGATCAATGCTTCAACTGGGCCAAGTGCGGAAGCTGAGCTTCAA 863
 DB 820 ACGGGGAAATACAAAGATCAATCTTTTAACTGGGCTAAAAATCCGCAAGTTGATTTAA 879
 QY 864 GAGGAAGCGCTTCTCATCAAGCTCCGGCCAGATGCGCAATAGTGCCTACAGATACCTT 923
 DB 880 GAGAAAGCATTTTCTCATCAAACTTCAATGCAATATCTTGGTGTGCAAGGATACCTT 939
 QY 924 GGAATTCGTATGCGAGTCGAGATTCTGCAAGTCTTCTGAAATAATCTGTGTGACCA 983

DB 940 GAGATTCACATGCGCAGCGAGATGCTCTGCAAGGCTTTCTGGAAGACTTGTGTGAATA 999
 QY 984 TCATGCTTCTTTAGACTTTTGAAGAGCCCAACCAAGCCCAAGCCCGTCTCTTTAG 1043
 DB 1000 CCATGCTTCTTCAAGCTTTTGGAGAGAGCCCAATCAAGCCCAAGCCCTACTCTGACAG 1059
 QY 1044 CCGGGGCTATCATATTTGGTTGAGTGTGAGCTCAGAGCAGGTTCTGACTATGTTAA 1103
 DB 1060 CAAGGTTTCAAGTTTCCGCTATATGATGAGACCAACCAAGGCACTTTTGGATATGGAG 1119
 QY 1104 AAGAGGAGACATTAAGAGATGAGCTTGAAGAGCA 1142
 DB 1120 AAAAGGAGGCTGAGAGCTTGCATTGAAAGAAACA 1158

Search completed: May 22, 2006, 01:14:34
 Job time : 1358.07 secs

This Page Blank (uspic)


```

gene
1. .475
/gene="DKFZp469C092"
CDS
100..3237
/gene="DKFZp469C092"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH91292.1"
/db_xref="GI:55729111"
/translation="MGEIRORPPGSRGLAPENSGISTLERGOKPPPPSGKLIVIKI
QMLDTQEPVEVPQAPKRVLDVNCNHLNVBGDFGLGECDDHKKITWDLILFVY
KQIRRKHVIVKVPKPPDHDTQLESLTRYLFALQVQDIAQGLKITNDLSALLI
SHIVSEIDFPDALREHLANKYIPQDADLEDKIVEHNHNIQGTPESPFLIEI
ARLEWYGIHLPAKDEGCTKINLAVANGTILVFOGTFTNFMNAKYKSLFKRFR
LTKRPDANSAYODTLEFLMASRPDKSPFKMCVEHHAFFRLEPKRPPKSLRSG
SEPSRGKQVLDVYKFGGHKQUPREKSKHISIRLALAPOLKVCYLPPLTF
ALTFREGASFEQSCQCKPKVSPGSGHPSVPFRPSAKNKQADGASATTEB
EEEVKDRQSKPQPPQSTSLTGSPLSLSVNSQGVAPANVTLSPLNSPOTLK
ASPLISPLNDQACPTDDEDEGRKRPPTDKAYFLAKSVTTERTYLQDLVITSWF
QSAVSEKIDEMPAKSLIPNPEPLHKFTPTNFKIEIQRALMEGRSNQIYDQRI
DVLNKOIGMKHLAVHLMGSEALENGISRSRLBNFCRDELOKVCYLPPLTF
LRLRHMHYKQVLERLCKHPPSHADPRCDRALAGITEMVQALHGTMIKMEFOKL
HELKDLIGTIDNLVYPRGEPFIRLGSLSKSGKLOQRMFLPNDLILTSRSLTASNO
FVHQQLPLYGMTIKESDEWVPHCLTLRGROSTIVASSRSEKWEVEDIQALID
LAEKNSLAPERFLASSPPDNKSPDEXTAADQSEEDLSASRTSLERQAFHRENTVHV
CMHRNTSVMDVFSVAVENQLSGNLIRKFNKSNQKLMVAVTFNFCLEFFYKSHQDNP
LASPLIGYSLTIPTESENIHKDYVKLHFKSHVYVYFRAESEYTFERMEVIRSATSS
ASRVVSSHESLKV"

```

ORIGIN

```

Query Match      97.0%; Score 3339.8; DB 6; Length 4775;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 3305; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

```

```

OY 1 CGCCGAGCCCGCGGCTGTGAGATATTTCTTAAGCCGCTTTCATCATGGAGAAATA 60
DB |||
OY 52 CGCCGAGCCCGCGGCTGTGAGATATTTCTTAAGCCATTCATCATGGAGAAATA 111
DB |||
OY 61 GAGCAGAGCCGAGCCCGAGATACAGACTGGGGGCCCCGAGAAATTCGGGGATCACTACC 120
DB |||
OY 112 GAGCAGAGCCGAGCCCGAGATACAGACTGGGGGCCCCGAGAAATTCGGGGATCACTACC 171
DB |||
OY 121 TTGGAACTGTGACAGAAAGCCGCCCAACACTTCAGAAAACTGCTGTTCATCAAAATC 180
DB |||
OY 172 TTGGAACTGTGACAGAAAGCCGCCCAACACTTCAGAAAACTGCTGTTCATCAAAATC 231
DB |||
OY 181 CAGATGCTGTGATACACCCAGAGGAGATTTGAAGTTCCAAAGACTCTCGGAGAGTG 240
DB |||
OY 232 CAGATGCTGTGATACACCCAGAGGAGATTTGAAGTTCCAAAGACTCTCGGAGAGTG 291
DB |||
OY 241 CTGCTGATGCAATTTGCAACCACTCAACCTGTGGAAGTGACTAATTTGGCTCGAG 300
DB |||
OY 292 CTGCTGATGCAATTTGCAACCACTCAACCTGTGGAAGTGACTAATTTGGCTCGAG 351
DB |||
OY 301 TTTCCTGATCACAAAAGATCAGGCTGTGGCTGGATCTCTTAAACCATTTGTAAACAG 360
DB |||
OY 352 TGCTCGATCACAAAAGATCAGGCTGTGGCTGGATCTCTTAAACCCTTGTGAACAG 411
DB |||
OY 361 ATTAAGAGCCCAAGAGACTGTTTGAAGTTGTGTGAAATTTCTTCGCTGACAC 420
DB |||
OY 412 ATCAGAGGCAAAAGATGTTGTTAAAGTTGTGTGAAATTTCTTCGCTGACAC 471
DB |||
OY 421 ACACAACCTCAAGAAACCTCAACAAGTACTGTGCGCTGAGAGTGAGAGCACTTG 480
DB |||
OY 472 ACACAACCTCAAGAAACCTCAACAAGTACTGTGCGCTGAGAGTGAGAGCACTTG 531
DB |||
OY 481 GCTCAAGGAGGTTGACGTGTAAAGACACAGGAGACTCTCTTGAATTTCAACATTTG 540
DB |||
OY 532 GCTCAAGGAGGTTGACGTGTAAAGACACAGGAGACTCTCTTGAATTTCAACATTTG 591
DB |||
OY 541 CAATCTGAATTTGGGATTTTGAAGCTTGGAGAGAGAGACTTACCAAAAATAA 600
DB |||
OY 592 CAATCTGAATTTGGGATTTTGAAGCTTGGAGAGAGAGACTTACCAAAAATAA 651
DB |||

```

```

OY 601 TACATACCTCAGACAGAGCACTAGAGACAAATCTGTGAATTTCAACATAACCAAT 660
DB |||
OY 652 TACATACCTCAGACAGAGCACTAGAGACAAATCTGTGAATTTCAACATAACCAAT 711
DB |||
OY 661 GGAACAAACACAGAGCAATCAGATTTCCAGCTCTTAGAGATTGCCCTGCTAGAGATG 720
DB |||
OY 712 GGAACAAACACAGAGCAATCAGATTTCCAGCTCTTAGAGATTGCCCTGCTAGAGATG 771
DB |||
OY 721 TATGAATCCGGTTGACCCCGGCAAGGACAGGAGAGGACAGAGATCAATCTGGCCGTT 780
DB |||
OY 772 TATGAATCCGGTTGACCCCGGCAAGGAGAGGAGGACAGAGATCAATCTGGCCGTT 831
DB |||
OY 781 GCCAACACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCC 840
DB |||
OY 832 GCCAACACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCC 891
DB |||
OY 841 AAGTTCGGAGCTGAGCTTCAAGAGAAAGCCCTTTCATCAAGTCCGCGCAGATGCC 900
DB |||
OY 892 AAGTTCGGAGCTGAGCTTCAAGAGAAAGCCCTTTCATCAAGTCCGCGCAGATGCC 951
DB |||
OY 901 AATAGGCGTACAGAGATACCTTGAATTCCTGATGGCCAGTCCGGGATTTTGCAAGTCC 960
DB |||
OY 1021 AAGCCCAAGCCGCTCTTTAGCCGGGGGTATCATTTCCGTTCAAGTGTGGACTCAG 1080
DB |||
OY 1072 AAGCCCAAGCCGCTCTTTAGCCGGGGGTATCATTTCCGTTCAAGTGTGGACTCAG 1131
DB |||
OY 1081 AAGCAGGTTCTGACATGTTTAAAGAGAGACATTAAGAGTGCAGTTTGAAGAAG 1140
DB |||
OY 1132 AAGCAGGTTCTGACATGTTTAAAGAGAGACATTAAGAGTGCAGTTTGAAGAAG 1191
DB |||
OY 1141 CACAGCAAGATTCATTCATTCAGAGCTTCTTCAAGACCTTACAGAACTGAATTCGAA 1200
DB |||
OY 1192 CACAGCAAGATTCATTCATTCAGAGCTTCTTCAAGACCTTACAGAACTGAATTCGAA 1251
DB |||
OY 1201 GTGCTGAGAGCTTCTCAGAGAGCACAGCTTACATTTGGAGAAAGTGGCCGAATTCGA 1260
DB |||
OY 1252 GTACTGAGAGCTTCTCAGAGAGCACAGCTTACATTTGGAGAAAGTGGCCGAATTCGA 1311
DB |||
OY 1261 GGGGGCAGAGCTGCGCGGAGAGAAAGCAAGAGTTTCCGCGGGGAGCCGGGGTGC 1320
DB |||
OY 1312 GGGGGCAGAGCTGCGCGGAGAGAAAGCCCGGGGTAAACAGAGCGGAGCGGCGC 1371
DB |||
OY 1321 CACCGAGGCTTGCAGCGGAGAGAAAGCCCGGGGTAAACAGAGCGGAGCGGCGC 1380
DB |||
OY 1372 CACCGAGGCTTGCAGCGGAGAGAAAGCCCGGGGTAAACAGAGCGGAGCGGCGC 1431
DB |||
OY 1381 TCGGGGCCACGAGAGAGAGAGAGAGTCTTTAAGATTAAGACCAAGAGATTAACCT 1440
DB |||
OY 1432 TCGGGGCCACGAGAGAGAGAGAGAGTCTTTAAGATTAAGACCAAGAGATTAACCT 1491
DB |||
OY 1441 CAGCCCCGACAGCAAGCAAGGCTCCCTGACTGGCAAGTCTCACTTTCCGAGCTGTCT 1500
DB |||
OY 1492 CAGCCCCGACAGCAAGCAAGGCTCCCTGAACCGGACGCTCACTTTCTGAGCTGTCT 1551
DB |||
OY 1501 GTGAACCTGACAGGGGAGAGTGCCCTTGCCAAGTGAACCTTGTCTTCCAACTGAGCCCC 1560
DB |||
OY 1552 GTGAACCTGACAGGGGAGAGTGCCCTTGCCAAGTGAACCTTGTCTTCCAACTGAGCCCC 1611
DB |||
OY 1561 GACACCAAGCAGGCTTCTTCCCTTGAATCAGCCCGCTGTGAATGAACAGGCTTGGCCCCGG 1620
DB |||
OY 1612 GACACCAAGCAGGCTTCTTCCCTTGAATCAGCCCGCTGTGAATGAACAGGCTTGGCCCCGG 1671
DB |||
OY 1621 ACAGAGATGAGAGTGGGGCCGAGAGAAAGATTTCCCAACTGATTAACGTAACCTTCATA 1680
DB |||
OY 1672 ACAGAGATGAGAGTGGGGCCGAGAGAAAGATTTCCCAACTGATTAACGTAACCTTCATA 1731
DB |||
OY 1681 GCTAAGGAATGTCTTACCAACGAGGAAATATCTGAAGGATCTCGAAGTTATCACTTGC 1740
DB |||

```

```

Db      1732 GCTAAGAGAGTGTCTACCAACGAGCGAACTATCTGAAGATCTCGAAGTATCACTTG 1791
Qy      1741 TGGTTTCAGAGCACTAGTACCAAGAGAGCGCAATGCGGAGCACTGAAAGTCTCAT 1800
Db      1792 TGGTTTCAGAGCGCACTAGAGCAAGAGAGCGCATGCGAAGACCTGAAAGTCTCAT 1851
Qy      1801 TTCGCAATTTTGAACCTTTGCACAATTTTCACTAATTTTCTCAAGGAAATTTAGCAA 1860
Db      1852 TTCGCAATTTTGAACCTTTGCACAATTTTCACTAATTTTCTCAAGGAAATTTAGCAA 1911
Qy      1861 CGACTTGCCCTGTGGAGAGCGCTCAATATGCCAATTCAGAGATTAACCAAGATTCGGC 1920
Db      1912 CGACTTGCCCTGTGGAGAGCGCTCAATATGCCAATTCAGAGATTAACCAAGATTCGGC 1971
Qy      1921 GATGTATGTGTAAGAACTATTCAGGGCATGAGCACTGCGGCTCACTGTGGAGCAC 1980
Db      1972 GATGTATGTGTAAGAACTATTCAGGGCATGAGCACTGCGGCTCACTGTGGAGCAC 2031
Qy      1981 AGCGAGGCGCTGGAGGCGCTGAGAAATGGAATGGAATGGAATGGAATGGAATGGAATG 2040
Db      2032 AGCGAGGCGCTGGAGGCGCTGAGAAATGGAATGGAATGGAATGGAATGGAATGGAATG 2091
Qy      2041 TGCAGAGACTTTGAGCTGAGAGAGGTGTGTTACTTACCGCTCAACAACCTTCTCTGCGG 2100
Db      2092 TGCAGAGACTTTGAGCTGAGAGAGGTGTGTTACTTACCGCTCAACAACCTTCTCTGCGG 2151
Qy      2101 CCACTGCAACCGGCTATGCACTACCAAGAGAGTCTTGAAGCGGCTGTGCAACAACCGG 2160
Db      2152 CCACTGCAACCGGCTATGCACTACCAAGAGAGTCTTGAAGCGGCTGTGCAACAACCGG 2211
Qy      2161 CGAGAGCGCGGCACTTCAAGGAGCTGCGGAGCGCTTTGGCAGAGATTCAGAGATGAGT 2220
Db      2212 CGAGAGCGCGGCACTTCAAGGAGCTGCGGAGCGCTTTGGCAGAGATTCAGAGATGAGT 2271
Qy      2221 GCACAGCTCCAGGATGCAATGATCAAGATGAGAAATTTCCAGAGCTGCAAGACTCAAG 2280
Db      2272 GCACAGCTCCAGGATGCAATGATCAAGATGAGAAATTTCCAGAGCTGCAAGACTCAAG 2331
Qy      2281 AAAAGATTTGATTTGATGCAATCTTTGTGTTCCGGGAAAGGAGTTTCACTCTGCGG 2340
Db      2332 AAAAGATTTGATTTGATGCAATCTTTGTGTTCCGGGAAAGGAGTTTCACTCTCTGCGG 2391
Qy      2341 AGCTTCAGAGAGCTCTCGGGGAAAGGAGCTCCAGAGCGCACTTCTCTCTGTTCAAGAC 2400
Db      2392 AGCTTCAGAGAGCTCTCTCGGGGAAAGGAGCTCCAGAGCGCACTTCTCTCTGTTCAAGAC 2451
Qy      2401 GTCTGCTATTAACGAGCGCGGAGCTGAGCGGCTCCCAATCACTTTAAAGTCCAGCGGAG 2460
Db      2452 GTCTGCTATTAACGAGCGCGGAGCTGAGCGGCTCCCAATCACTTTAAAGTCCAGCGGAG 2511
Qy      2461 GTCTGCTATTAAGCATGATGAGAGAGCGAAAGCATGAGGAGGAGTCCCACTGC 2520
Db      2512 GTCTGCTATTAAGCATGATGAGAGAGCGAAAGCATGAGGAGGAGTCCCACTGC 2571
Qy      2521 CTGAGCCCTCGGGGAGCGGAGCTGATCACTATGAGCGGCGGCTCTGCTGAGAGT 2580
Db      2572 CTGAGCCCTCGGGGAGCGGAGCTGATCACTATGAGCGGCGGCTCTGCTGAGAGT 2631
Qy      2581 GAGAAAGTGGGTTGAGAGCATTCAGATGAGCATTTGAGCTGCGGAGGAGAGAGCAGACCCC 2640
Db      2632 GAGAAAGTGGGTTGAGAGCATTCAGATGAGCATTTGAGCTGCGGAGGAGAGAGCAGACCCC 2691
Qy      2641 GCGCCCTGAGTCTTGCGGAGAGCGGCGGCTGAGCAACAAGTCCCTGATGAGAGCAGCGG 2700
Db      2692 GCGCCCTGAGTCTTGCGGAGAGCGGCGGCTGAGCAACAAGTCCCTGATGAGAGCAGCGG 2751
Qy      2701 GCGTACCCAGAGATGAGAGTATGAGCGCTGCGGAGCATGCGTGAAGCGGAGCGG 2760
Db      2752 GCGTACCCAGAGATGAGAGTATGAGCGCTGCGGAGCATGCGTGAAGCGGAGCGG 2811
Qy      2761 CCGCAGCGGAGCAACAATGATGAGCGTGTGCTGAGCAGCGCAACAAGCGTCTCATG 2820

```

```

Db      2812 CCGCAGCGGAGCAACAATGATGATGTGTGTGAGCAGCGCAACACAGCGTCTCATG 2871
Qy      2821 GTGAGCTTACGATTCGCAATGAGAAATCACTTTGTGAGAAACCTGCTGAGAAATTCAPA 2880
Db      2872 GTGAGCTTACGCTTTCAGATGAGAAATCACTTTGTGAGAAACCTGCTGAGAAATTCAPA 2931
Qy      2881 AACAGCAAGGAGGAGAGAGCTGTGGTGTGTTTCAAACTTGTGCTGTTCTTAC 2940
Db      2932 AACAGCAAGGAGGAGAGAGCTGTGGTGTGTTTCAAACTTGTGCTGTTCTTAC 2991
Qy      2941 AAATCAACACAGGACATCAATCCCTTGTGCAAGCTGCTGTGCTGCTGCTGCTGCTGCTG 3000
Db      2992 AAGTCAACACAGGACATCAATCCCTGCGCAGCTGCTGTGCTGCTGCTGCTGCTGCTGCTG 3051
Qy      3001 ATCCCTCTGAGTCCGAGAACTCCAGAAAGATAGTGTTCAGAGCTGCACTTCAAGTCC 3060
Db      3052 ATCCCTCTGAGTCCGAGAACTCCAGAAAGATAGTGTTCAGAGCTGCACTTCAAGTCC 3111
Qy      3061 CAGCTCTACTACTTTCAGGAGGAGAGAGAGTACAGCTTGCAGAAAGTGTGAGTATGATC 3120
Db      3112 CAGCTCTACTACTTTCAGGAGGAGAGAGTACAGCTTGCAGAAAGTGTGAGTATGATC 3171
Qy      3121 CCGAGTGCACCAAGCTCTGCTGCGAGCCCAAGTGTGAGCCCAAGAGTCTTGTG 3180
Db      3172 CCGAGGCGCACCAAGCTCTGCTGCGAGTCAAGTGTGAGCCCAAGAGTCTTGTG 3231
Qy      3181 TATTGATGCGCGGAGCACTCTGTTTCCGAGTGTGCTTCTGGAAGACGTTTCTT 3240
Db      3232 TATTGATGCGCGGAGCACTCTGTTTCCGAGTGTGCTTCTGGAAGACGTTTCTT 3291
Qy      3241 CTCTGATTTAAATGAGAGCTGTGTTAAATTAACAGCTGTGTTAAATCAAAATCATGAGCTT 3300
Db      3292 CTCTGATTTAAATGAGAGCTGTGTTAAATTAACAGCTGTGTTAAATCAAAATCATGAGCTT 3351
Qy      3301 CCCAGAGCTCTCTGCTGCTGCAAGCGCGG-TTTTAAATCCCGAGCTCTGAGGTTTGA 3359
Db      3352 CCCAGAGCTCTCTGCTGCTGCAAGCGCGGTTTAAATCCCGAGCTCTGAGGTTTGA 3411
Qy      3360 ATGAAGAGGCTTCCCACTTCCAGTCTGAGATCCGCTGGGGGCGCTTCTTAACTAGT 3419
Db      3412 ACMAAGAGGCTTCCCACTTCCAGTCTGAGATCCGCTGGGGGCGCTTCTTAACTAGT 3471
Qy      3420 GCCAGTATTTAAATCATTTGTCATT 3442
Db      3472 GCCAGTATTTAAATCATTTGTCATT 3494

RESULT 2
AY406167 3138 bp DNA linear GSS 15-DEC-2003
LOCUS Homo sapiens PARP1 gene, VIRIDAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY406167
VERSION AY406167.1 GI:39762141
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 3138)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3138)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,

```

Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 Location/Qualifiers
 source
 1. 3138
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>3138
 /gene="FARP1"
 /locus_tag="HCM2467"

ORIGIN

Query Match	91.1%;	Score 3136.4;	DB 14;	Length 3138;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 3137;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	49	ATGGGGAATAGAGAGAGGCGGACCCAGATCAGACTGGGGGCGCCGGAAATTGG	108	
DB	1	ATGGGGAATAGAGAGAGGCGGACCCAGATCAGACTGGGGGCGCCGGAAATTGG	60	
QY	109	GGGATCAGTACCTTGGAGCGTGACAGAACCGCCCAACACTTCAGAGAAACTCGTG	168	
DB	61	GGGATCAGTACCTTGGAGCGTGACAGAACCGCCCAACACTTCAGAGAAACTCGTG	120	
QY	169	TCCATCAAAATCCAGATGCTGATGACCCAGAGAGCATTTGAAGTTCCACAAAGACT	228	
DB	121	TCCATCAAAATCCAGATGCTGATGACCCAGAGAGCATTTGAAGTTCCACAAAGACT	180	
QY	229	CCGGGAAGGTGCTGTGATGATGCACTTCAACACTCGTGAAGGTGACTAT	288	
DB	181	CCGGGAAGGTGCTGTGATGATGCACTTCAACACTCGTGAAGGTGACTAT	240	
QY	289	TTTGCCCTGAGTTCTCTGATCACAAAAGATCAGCGTGTGGCTGATCTCTTAAACC	348	
DB	241	TTTGCCCTGAGTTCTCTGATCACAAAAGATCAGCGTGTGGCTGATCTCTTAAACC	300	
QY	349	ATTGTGAAACAGATTGAAAGGCCAAAGCATGTTGTTGTTAAGTTTGTGTAATCTTT	408	
DB	301	ATTGTGAAACAGATTGAAAGGCCAAAGCATGTTGTTGTTAAGTTTGTGTAATCTTT	360	
QY	409	CGGCTGACACACACAACTCCAGAGAACTCAGAAAGTACCTGTTGCGGCTGACGATG	468	
DB	361	CGGCTGACACACACAACTCCAGAGAACTCAGAAAGTACCTGTTGCGGCTGACGATG	420	
QY	469	AAGCAGACTTGGCTCAAGCAGGTGACGTGTATGACACACGCGCAGCTCTTGTATT	528	
DB	421	AAGCAGACTTGGCTCAAGCAGGTGACGTGTATGACACACGCGCAGCTCTTGTATT	480	
QY	529	TACACATTGTGCAATCTGAGATTGGGGATTTTGAAGCCTTGGACAGAGACTTAA	588	
DB	481	TACACATTGTGCAATCTGAGATTGGGGATTTTGAAGCCTTGGACAGAGACTTAA	540	
QY	589	GCAAAATAAATACTACCTCAGCAAGACGCACTAGAGACAAATACTGTGAAATTTCAC	648	
DB	541	GCAAAATAAATACTACCTCAGCAAGACGCACTAGAGACAAATACTGTGAAATTTCAC	600	
QY	649	CATAACCACTTGGACAAACACCAAGCAGATCAGATTTTCAGCTCTTAGAGATTGCCGT	708	
DB	601	CATAACCACTTGGACAAACACCAAGCAGATCAGATTTTCAGCTCTTAGAGATTGCCGT	660	
QY	709	CGGCTAGAGATGTATGGAATTCGGGTGACCCGGCCAAAGACAGGGAAGGCAGAAATC	768	
DB	661	CGGCTAGAGATGTATGGAATTCGGGTGACCCGGCCAAAGACAGGGAAGGCAGAAATC	720	
QY	769	AATTCGCGCTTGGCCAAACGAGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC	828	
DB	721	AATTCGCGCTTGGCCAAACGAGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC	780	
QY	829	TTCAACTGGGCCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCCTTCTCATCAAGCTTC	888	

DB	781	TTCAACTGGGCCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCCTTCTCATCAAGCTTC	840	
QY	889	CGGCGAGATGCCAATAATAGTCCTTACAGAGATACCTTGGAAATTCCTGATGGCCGATCGGGAT	948	
DB	841	CGGCGAGATGCCAATAATAGTCCTTACAGAGATACCTTGGAAATTCCTGATGGCCGATCGGGAT	900	
QY	949	TTCTGCAAGTCTTCTTGGAAATCTGTGTGAACATCATGCTCTTGTAGACTTTTGAAT	1008	
DB	901	TTCTGCAAGTCTTCTTGGAAATCTGTGTGAACATCATGCTCTTGTAGACTTTTGAAT	960	
QY	1009	GAGCCCAACCAAGCCCAAGCCGCTCTTTCAGCCGGGGTCAATTCGTTGATTCAGT	1068	
DB	961	GAGCCCAACCAAGCCCAAGCCGCTCTTTCAGCCGGGGTCAATTCGTTGATTCAGT	1020	
QY	1069	GGTGGGACTCAAGACAGTCTCGACTATGTATTAAGAGAGACATAAGAAAGGTGACG	1128	
DB	1021	GGTGGGACTCAAGACAGTCTCGACTATGTATTAAGAGAGACATAAGAAAGGTGACG	1080	
QY	1129	TTTGAAGGAAGCAGCAGAAATTCATTCTATCCGAGGCTTGTTCACAGCCTACAGAA	1188	
DB	1081	TTTGAAGGAAGCAGCAGAAATTCATTCTATCCGAGGCTTGTTCACAGCCTACAGAA	1140	
QY	1189	CTGAATTGGAAGTGTGAGAGAGTCTCAGACAGAGCACCAGCTTACATTTGAGAAAGT	1248	
DB	1141	CTGAATTGGAAGTGTGAGAGAGTCTCAGACAGAGCACCAGCTTACATTTGAGAAAGT	1200	
QY	1249	GCGGAATCTCCAGGGGGGCGCAGAGCTGCGGCGAGAAAGAACCAAGGTTTCCGCGGG	1308	
DB	1201	GCGGAATCTCCAGGGGGGCGCAGAGCTGCGGCGAGAAAGAACCAAGGTTTCCGCGGG	1260	
QY	1309	GAGCCGGGGTCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGTTTAAACAAGAGCG	1368	
DB	1261	GAGCCGGGGTCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGTTTAAACAAGAGCG	1320	
QY	1369	GAGGAGCCGCTCGGCGCCGACGAGAGAGAGAGAGTCTTAAAGATAGAACCCAG	1428	
DB	1321	GAGGAGCCGCTCGGCGCCGACGAGAGAGAGAGAGTCTTAAAGATAGAACCCAG	1380	
QY	1429	CAGAGTAAACCTCAGCCCGGAGCAGAGCAGAGCTCTGATCGGCGAGTCTCAACCTT	1488	
DB	1381	CAGAGTAAACCTCAGCCCGGAGCAGAGCAGAGCTCTGATCGGCGAGTCTCAACCTT	1440	
QY	1489	TCGAGCTGTCTGTAACTCCGAGGGGGAGTGGCCCTGCGCAAGTGAATGTTCTCC	1548	
DB	1441	TCGAGCTGTCTGTAACTCCGAGGGGGAGTGGCCCTGCGCAAGTGAATGTTCTCC	1500	
QY	1549	AACTGAGCCCGGACACAGCAGGCTCTCTCTTGATCAGCCGCTGCTGAATGACAG	1608	
DB	1501	AACTGAGCCCGGACACAGCAGGCTCTCTCTTGATCAGCCGCTGCTGAATGACAG	1560	
QY	1609	GCTGCCCCCGGAGGAGATGAGAGTGAAGGCGGAGGAAGATTCCTCAACTGATAA	1668	
DB	1561	GCTGCCCCCGGAGGAGATGAGAGTGAAGGCGGAGGAAGATTCCTCAACTGATAA	1620	
QY	1669	GCGTACTTCAATGACTAAGGAAGTGTCTACACAGAGCAATATCTGAAGATCTCGAA	1728	
DB	1621	GCGTACTTCAATGACTAAGGAAGTGTCTACACAGAGCAATATCTGAAGATCTCGAA	1680	
QY	1729	GTTATCACTTGGTGTTCAGAGCAAGTGAACAAAGAGACGCAATGCCGAGACACTG	1788	
DB	1681	GTTATCACTTGGTGTTCAGAGCAAGTGAACAAAGAGACGCAATGCCGAGACACTG	1740	
QY	1789	AAAAGTCTATTTCCCGAATTTTGAACCTTGGCAAAATTCATCTAATTTTTCGAAG	1848	
DB	1741	AAAAGTCTATTTCCCGAATTTTGAACCTTGGCAAAATTCATCTAATTTTTCGAAG	1800	
QY	1849	GAAATTGAGCAACGACTTGGCCCTGTGGGAAGGCGCTCAATAGCCCAATTCAGAGATTAC	1908	
DB	1801	GAAATTGAGCAACGACTTGGCCCTGTGGGAAGGCGCTCAATAGCCCAATTCAGAGATTAC	1860	
QY	1909	CAGAAATGGGAGATGATAGCTTGAAGAACATTCAGGGCATGAAGCACTTGGCGCTCAC	1968	

Db 1861 CAAGAATCGGCGATGTCATGTAAGAACATTCAAGGCGATGAAGCACTGGGGCTCAC 1920

Qy 1969 CTGTGGAAGCAGACGAGGCTTGGAGCCCTTGGAGAAATGGAATCAAGAGCTCCGGCGG 2028

Db 1921 CTGTGGAAGCAGACGAGGCTTGGAGGCCCTGGAAGATGAATCAAGAGCTCCGGCGG 1980

Qy 2029 CTGGAAGAACTTTCAGAGAGACTTGGAGCTCAGAGAGGTGTGTTACCTACCGCTCAACACC 2088

Db 1981 CTGGAAGAACTTTCAGAGAGACTTGGAGCTCAGAGAGGTGTGTTACCTACCGCTCAACACC 2040

Qy 2089 TTCTCTCTGAGGCACTGCAACGGCTCATGCACTAACAAGAGTCTTGAAGCGGCTGTGC 2148

Db 2041 TTCTCTCTGAGGCACTGCAACGGCTCATGCACTAACAAGAGTCTTGAAGCGGCTGTGC 2100

Qy 2149 AAACACACACCCCGGAGCCAGCCGCACTTCAAGGAGCTGCGGAGCCGTTTGGCAGAGATC 2208

Db 2101 AAACACACACCCCGGAGCCAGCCGCACTTCAAGGAGCTGCGGAGCCGTTTGGCAGAGATC 2160

Qy 2209 ACGGAGATGGTGGCAGACGCTCCAGGTAAGATGATCAAGATGGAATTTCCAGAAAGCTG 2268

Db 2161 ACGGAGATGGTGGCAGACGCTCCAGGTAAGATGATCAAGATGGAATTTCCAGAAAGCTG 2220

Qy 2269 CACGAACTCAAGAAAGATTGATTGGCAATTGCAATCTTGTGTTCCGGGAGAGAGTTTC 2328

Db 2221 CACGAACTCAAGAAAGATTGATTGGCAATTGCAATCTTGTGTTCCGGGAGAGAGTTTC 2280

Qy 2329 ATCCGCTTGGGCGCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAGCGCATGTTCTTC 2388

Db 2281 ATCCGCTTGGGCGCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAGCGCATGTTCTTC 2340

Qy 2389 CTGTTCACGACGCTCTGCTATACAGAGCGGGGGCTGACGGGCTCCAAATCAGTTTAAA 2448

Db 2341 CTGTTCACGACGCTCTGCTATACAGAGCGGGGGCTGACGGGCTCCAAATCAGTTTAAA 2400

Qy 2449 GTTCACGGGCGAGCTCCCGCTCTATAGGATGCAATTTGAGAGAGCGAAGACGAGTGGGG 2508

Db 2401 GTTCACGGGCGAGCTCCCGCTCTATAGGATGCAATTTGAGAGAGCGAAGACGAGTGGGG 2460

Qy 2509 GTGCCCCAATGCTGACCTCTCGGGGCGAGCGGCACTCATCTGTGGCGGCAATGTTCT 2568

Db 2461 GTGCCCCAATGCTGACCTCTCGGGGCGAGCGGCACTCATCTGTGGCGGCAATGTTCT 2520

Qy 2569 CGGTCCGAGATGAGAGAGTGGTGTGAGAGATCAATGAGTGGCAATTTGACCTGGCGAGAA 2628

Db 2521 CGGTCCGAGATGAGAGAGTGGTGTGAGAGATCAATGAGTGGCAATTTGACCTGGCGAGAA 2580

Qy 2629 AGCAGACGCCCCCGCTGAGTTCCTGGCCAGAGCCCCCTGACAAACAAGTCCCTGAT 2688

Db 2581 AGCAGACGCCCCCGCTGAGTTCCTGGCCAGAGCCCCCTGACAAACAAGTCCCTGAT 2640

Qy 2689 GAAGCCACCGCGGCTGACAGAGATGAGAGTGAAGTGGCGCTTCCGCGCATGCTG 2748

Db 2641 GAAGCCACCGCGGCTGACAGAGATGAGAGTGAAGTGGCGCTTCCGCGCATGCTG 2700

Qy 2749 GAGCGCAGGCCCCCGACCGCGGCAACAAATGATGACGTGTGCTGGCAGCGCAACCC 2808

Db 2701 GAGCGCAGGCCCCCGACCGCGGCAACAAATGATGACGTGTGCTGGCAGCGCAACCC 2760

Qy 2809 AGCGTCTCAATGATGATTCAGCATTCAGAGTGAAGATCAATGATTCCTGAAAACCTGCTG 2868

Db 2761 AGCGTCTCAATGATGATTCAGCATTCAGAGTGAAGATCAATGATTCCTGAAAACCTGCTG 2820

Qy 2869 AGGAATTTCAAAAACAGCAACGGGTGGCAGAACTGTGGGTGTTCACAAACTTCTGC 2928

Db 2821 AGGAATTTCAAAAACAGCAACGGGTGGCAGAACTGTGGGTGTTCACAAACTTCTGC 2880

Qy 2929 CTGTTCCTTACAAATACACACAGAGCAATCATCCCTTGCAGAGCTGCTGTCTGCGG 2988

Db 2881 CTGTTCCTTACAAATACACACAGAGCAATCATCCCTTGCAGAGCTGCTGTCTGCGG 2940

Qy 2989 TACTGCTCAACATCCCTCTGTAGTCCGAGAACATCCAGAAAGATCACTGTTCAGAGCTG 3048

Db 2941 TACTGCTCAACATCCCTCTGTAGTCCGAGAACATCCAGAAAGATCACTGTTCAGAGCTG 3000

Qy 3049 CACTTCAAGTCCCAAGTCTACTACTTCAAGGGGGAAGGAGTACAGCTTCAAGAAAGTGG 3108

Db 3001 CACTTCAAGTCCCAAGTCTACTACTTCAAGGGGGAAGGAGTACAGCTTCAAGAAAGTGG 3060

Qy 3109 ATGGAAGTATCCGACAGTCCACACAGCTGCTCCGCAACCCCACTGTGTGAGCCACAA 3168

Db 3061 ATGGAAGTATCCGACAGTCCACACAGCTGCTCCGCAACCCCACTGTGTGAGTACAA 3120

Qy 3169 GAGTCTCTGTGTATGA 3186

Db 3121 GAGTCTCTGTGTATGA 3138

RESULT 3
AY406168 3138 bp DNA linear GSS 15-DEC-2003
LOCUS
DEFINITION Pan troglodytes FARP1 gene, VIRUTAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY406168
ACCESSION AY406168.1 GI:39762142
VERSION
KEYWORDS
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Pan.
1 (bases 1 to 3138)
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Snihsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL
PUBMED 14671302
REFERENCE 2 (bases 1 to 3138)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Snihsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
TITLE
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..3138
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>3138
/gene="FARP1"
/locus_tag="HCM2467"
ORIGIN
Query Match 86.2%; Score 2968.2; DB 14; Length 3138;
Best Local Similarity 94.8%; Pred.No. 0;
Matches 2976; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
Qy 49 ATGGAGAAATAGAGCAGAGGCGGACCCCAAGATACAGACTGGGGCCCCGGAAAAATTGG 108
Db 1 ATGGAGAAATAGAGCAGAGGCGGACCCCAAGATGCGACTGGGGCCCCGGAAAAATTGG 60
Qy 109 GGGATCAGTACTTGGAACTGGGACAGAAAGCCGCCCAACCTTTCAGAAAACTCGTG 168
Db 61 GGGATCAGTACTTGGAACTGGGACAGAAAGCCGCCCAACCTTTCAGAAAACTCGTG 120
Qy 169 TGCATCAAAAATCAATATGATGATGACACCCAGAGAGCAATTTGAAGTTCACAAAGAGCT 228
Db 121 TGCATCAAAAATCAATATGATGATGACACCCAGAGAGCAATTTGAAGTTCACANNNNNNNN 180
Qy 229 CCTGGGAAGTCTGCTGATCAATTTGCAACCACTCAACCTCGTGGAGGTGACTAT 288

[illegible]

Dd	1261	GAGCCGGGGGTGCAACCCGAGCCCTCTGGCCGACGAGGAAGACC	CCGCGGGTAAACMAAGCAGGC	1320
Qy	1369	GACGAGACCCGCTTCGCGCGCCCAACGAGGAAGAGAGGTCGTTAAAGATAGGACCAG		1428
Dd	1321	GACGAGACCCGCTTCGCGCGCCCAACGAGGAAGAGAGAGGTCGTTAAAGATAGGACCAG		1380
Qy	1429	CAGAGTAAACCTCAGACCCCGCGAGCCCAAGACAGGGCTCCCTGACTGGCACTCCCACTT		1488
Dd	1381	CAGGTAAACCTCAGACCCCGCGAGCCCAAGACAGGGCTCCCTGACTGGCACTCCCACTT		1440
Qy	1489	TCCGAGCTGTCTGTGAACCTCGCAGGGGGAGTGGCCCTTCCCAACGAGCTTGTCTCC		1548
Dd	1441	TCCGAGCTGTCTGTGAACCTCGCAGGGGGAGTGGCCCTTCCCAACGAGCTTGTCTCC		1500
Qy	1549	AACTGAGCCCCGACACCAAGCAGCCCTCTCCCTGATCAAGCCCGCTGTAATGACAG		1608
Dd	1501	AACTGAGCCCCGACACCAAGCAGCCCTCTCCCTGATCAAGCCCGCTGTAATGACAG		1560
Qy	1609	GCTGCCCCCGGAGCAGACGATGAGAGTGAAGGGCCGAGAGAAAGATTCCCACTGATAAA		1668
Dd	1561	GCTGCCCCCGGAGCAGACGATGAGAGTGAAGGGCCGAGAGAAAGATTCCCACTGATAAA		1620
Qy	1669	GCGTACTTCATAGCTAAGAAAGTGTCTAACCGAGCGAAATATCTGANGATCTCGAA		1728
Dd	1621	GCGTACTTCATAGCTAAGAAAGTGTCTAACCGAGCGAAATATCTGANGATCTTGAA		1680
Qy	1729	GTTATCACTTCGAGGTTTCAGAGCAGATGAGCAAAAGAGACGCAATCGCGAAGCAGTG		1788
Dd	1681	GTTATCACTTCGAGGTTTCAGAGCAGATGAGCAAAAGAGACGCAATCGCGAAGCAGTG		1740
Qy	1789	AAAAGTCTCATATTTCCGAAATTTTGAACCTTTGCACAAATTTCACTAAATTTTCTCAG		1848
Dd	1741	AAAAGTCTCATATTTCCGAAATTTTGAACCTTTGCACAAATTTCACTAAATTTTCTCAG		1800
Qy	1849	GAAATTAAGCAACGACTTGGCCCTGTGGGAAGGCGCGTCAAAATCCCAATCAGAGATTAC		1908
Dd	1801	GAAATTAAGCAACGACTTGGCCCTGTGGGAAGGCGCGTCAAAATCCCAATCAGAGATTAC		1860
Qy	1909	CAAAAGATTCGGCGCATGTCTATGCTGGAAGAACATTTACAGGGCATGAGACCTTGGCGGCTCAC		1968
Dd	1861	CAAAAGATTCGGCGCATGTCTATGCTGGAAGAACATTTACAGGGCATGAGACCTTGGCGGCTCAC		1920
Qy	1969	CTGTGGAAGCAGACGAGGCGCTTGGAGGCCCTTGGAGAAATGAAATCAAGACTCCCGCGG		2028
Dd	1921	CTGTGGAAGCAGACGAGGCGCTTGGAGGCCCTTGGAGAAATGAAATCAAGACTCCCGCGG		1980
Qy	2029	CTGAGAAACTTCTGACAGACTTTGAGCTCAGAGAGTGTGTTACTCAACGCTCAACACC		2088
Dd	1981	CTGAGAAACTTCTGACAGACTTTGAGCTCAGAGAGTGTGTTACTCAACGCTCAACACC		2040
Qy	2089	TTCTCTCTCGGCGCACTGCAACGCGCTCATGCACTACAAGCAGGTCTTGAGCGGCTGTGC		2148
Dd	2041	TTCTCTCTCGGCGCACTGCAACGCGCTCATGCACTACAAGCAGGTCTTGAGCGGCTGTGC		2100
Qy	2149	AAAACACACCCCGCGAGACCAAGCGCACTTCAAGGGACCTGCCAGACCGCTTTGGCAGAGATC		2208
Dd	2101	AAAACACACCCCGCGAGANN		2160
Qy	2209	ACGAGAGATGTGTGCAACGCTCCACGATACGATGATCAAGATGAGAAATTTTCAAGACTG		2268
Dd	2161	ACGAGAGATGTGTGCAACGCTCCACGATACGATGATCAAGATGAGAAATTTTCAAGACTG		2220
Qy	2269	CACGAACTCAAGAAAGATTGATTGGATTTGACAAATTTTGTGGTTCCGGGAAGGGAATTC		2328
Dd	2221	CACGAACTCAAGAAAGATTGATTGGATTTGACAAATTTTGTGGTTCCGGGAAGGGAATTC		2280
Qy	2329	ATTCGCTGTGGGACGCTCAGCAAGCTCTCGGGGAAGGGGCTCCAGCAGCCATGTTCTTC		2388
Dd	2281	ATTCGCTGTGGGACGCTCAGCAAGCTCTCGGGGAAGGGGCTCCAGCAGCCATGTTCTTC		2340
Qy	2389	CTGTTCAACGACGTCTGTCTATAACGAGCGCGGGGCTGACGGCTTCAATACGATTAAA		2448
Dd	2341	CTGTTCAACGACGTCTGTCTATAACGAGCGCGGGGCTGACGGCTTCAATACGATTAAA		2400

QY 2449 GTCCAGGGGAGCTCCCGCTCTATGGCATGCAATTGAGGAGAGCCGAAGACGAGTGGGG 2508
 DB 2401 GTCCAGGGGAGCTCCCGCTCTATGGCATGCAATTGAGGAGAGCCGAAGACGAGTGGGG 2460
 QY 2509 GTGCCCCACTGCTGAGCCCTCCGGGGCCAGCGGAGTCAATCATCTGAGCGCCAGTTCT 2568
 DB 2461 GTGCCCCACTGCTGAGCCCTCCGGGGCCAGCGGAGTCAATCATCTGAGCGCCAGTTCT 2520
 QY 2569 CGGTCCGAGATGAGAAAGTGGGTGAGAGACATCCAGATGGCCATTGACTCTGGGGAGAG 2628
 DB 2521 CGGTCCGAGATGAGAAAGTGGGTGAGAGACATCCAGATGGCCATTGACTCTGGGGAGAG 2580
 QY 2629 AGCAGAGCCCGCCCGCTGAGTTCTTGGCCAGAGCCCGCTGACAAAGTCCCGTAT 2688
 DB 2581 AGCAGAGCCCGCCCGCTGAGTTCTTGGCCAGAGCCCGCTGACAAAGTCCCGTAT 2640
 QY 2689 GAAGCCACCGCGGCTGACAGAGTCAAGAGATGACTGAGCGCTTGGCCAGATCGCTG 2748
 DB 2641 GAAGCCACCGCGGCTGACAGAGTCAAGAGATGACTGAGCGCTTGGCCAGATCGCTG 2700
 QY 2749 GAAGCCACCGCGGCTGACAGAGTCAAGAGATGACTGAGCGCTTGGCCAGATCGCTG 2808
 DB 2701 GAAGCCACCGCGGCTGACAGAGTCAAGAGATGACTGAGCGCTTGGCCAGATCGCTG 2760
 QY 2809 AGGCTCTCATGTGAGCTTCAAGATGCAAGTGAAGATCATGTTGCTGAAAACCTGCTG 2868
 DB 2761 AGGCTCTCATGTGAGCTTCAAGATGCAAGTGAAGATCATGTTGCTGAAAACCTGCTG 2820
 QY 2869 AGGAAATTCAAAAACAGCAACGGGTGGCAGAAAGCTGTGGTGTGTTCACAAATCTTGC 2928
 DB 2821 AGGAAATTCAAAAACAGCAACGGGTGGCAGAAAGCTGTGGTGTGTTCACAAATCTTGC 2880
 QY 2929 CTGTCTTCTTCAAAATACACACAGACATCATCTCTTCCAGCTGCTTCTGCTGGC 2988
 DB 2881 CTGTCTTCTTCAAAATACACACAGACATCATCTCTTCCAGCTGCTTCTGCTGGC 2940
 QY 2989 TACTCGCTACATCCCTCTGAGTCCGAGAACATCCAGAAAGATCAAGTTCAGACTG 3048
 DB 2941 TACTCGCTACATCCCTCTGAGTCCGAGAACATCCAGAAAGATCAAGTTCAGACTG 3000
 QY 3049 CACTTCAGTCCGACGCTCTACTACTTTCAGGGCGGAAAGCAGATGACATCTTGAAGGTG 3108
 DB 3001 CACTTCAGTCCGACGCTCTACTACTTTCAGGGCGGAAAGCAGATGACATCTTGAAGGTG 3060
 QY 3109 ATGGAAGTATCCGCACTGCTGCTTCCGAGCCCACTGTTGAGCCACAA 3168
 DB 3061 ATGGAAGTATCCGCACTGCTGCTTCCGAGCCCACTGTTGAGCCACAA 3120
 QY 3169 GAGTCTCTTGTATTGA 3186
 DB 3121 GAGTCTCTTGTATTGA 3138

RESULT 4
 AY406169 3147 bp DNA linear GSS 15-DEC-2003
 LOCUS AY406169
 DEFINITION Mus musculus PARP1 gene, VIRUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 VERSION AY406169
 KEYWORDS AY406169.1 GI:39762143
 SOURCE GSS.
 ORGANISM Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murinae; Mus.
 1 (bases 1 to 3147)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous

JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 3147)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1..3147
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>3147
 /gene="PARP1"
 /locus_tag="HMC2467"
 ORIGIN
 Query Match 70.5%; Score 2427.4; DB 14; Length 3147;
 Best Local Similarity 86.2%; Pred. No. 0;
 Matches 2712; Conservative 0; Mismatches 426; Indels 9; Gaps 2;
 QY 49 ATGGGAAATATGAGCAGAGCGCCAGCCAGATCCAGACTGGGGCCCCGGAAATTCG 108
 DB 1 ATGGGAAATATGAGCAGAGCGCCAGCCAGATCCAGACTGGGGCCCCGGAAATTCG 60
 QY 109 GGGATAGTACTCTGGAAGTGGACAGAGCGCGCCCAACCTTCAGAAAACCTCGT 168
 DB 61 GGGATAGTACTCTGGAAGTGGACAGAGCGCGCCCAACCTTCAGAAAACCTTATG 120
 QY 169 TCCATCAAAATCCAGATGCTGATGACACCCAGAGGCAATTTGAAGTTCACAAAGACT 228
 DB 121 ACTGTCAAAATCCAGATGCTGATGACACCCAGAGGCAATTTGAAGTTCACAAAGAGG 180
 QY 229 CCGGGAAGTGTGCTGATGCAATTTGCAACCACTTAACCTGCTGGAAGTACTAT 288
 DB 181 CCGGGAAGTGTGCTGATGCAATTTGCAACCACTTAACCTGCTGGAAGGCACTATC 240
 QY 289 TTGGGCTGAGTTTCTGATCAAAAAGATCAGGGTGGCTGGATCTCTTAAACC 348
 DB 241 TTGGGCTGAGTTTCTGATCAAAAAGATCAGGGTGGCTGGATCTCTTAAAGCCC 300
 QY 349 ATTGTGAAACAGATTAGAAAGCCAAAGCAGTGTGTTAGTGTGTGTAATTCCTT 408
 DB 301 ATTGTGAAACAGATTAGAAAGCCAAAGCAGTGTGTTAGTGTGTGTAATTCCTT 360
 QY 409 CCGCTGACACACACATCCCAAGAAACTCAAGATGACTGTTGGCGCTGCAGGTG 468
 DB 361 CCGCTGACACACACATCCCAAGAAACTCAAGATGACTGTTGGCGCTGCAGGTG 420
 QY 469 AAGCAGACTTGGCTCAAGGAGTTGAGTGTATGACACAGGCACTGCTTGATT 528
 DB 421 AAGCAGACTTGGCTCAAGGAGTTGAGTGTATGACACAGGCACTGCTTGATT 480
 QY 529 TCACATTTGTGCAATCTGAGATTGGGATTTTGTGTAAGCTTGGACAGAGCACTTA 588
 DB 481 TCACATTTGTGCAATCTGAGATTGGGATTTTGTGTAAGCTTGGACAGAGCACTTA 540
 QY 589 GCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 648
 DB 541 GCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
 QY 649 CATTAACCAATTGGAACAACACAGAGAAATGAGATTTTCAGTCTCTAGAGTGGCGGT 708
 DB 601 CACAGCCAGTTGAGCAACCCCGGAGAGTGAATTTTCAGTCTCTAGAGTGGCGGT 660
 QY 709 CGGCTAGAGATGTATGAAATCCGTTTGCACCGGCAAGAGCAGGAAAGCAGAGATC 768

Db 661 CGGCTAGAGATGTACGGAAATCCGGTTACACCTCGCCAAAGACAGGAGGCAACAAGATC 720
Qy 769 AATTCGGCCGTTGCCACAACGGGAATTCTAGTGTTCAGGGTTTCACTAAGATCAATGCG 828
Db 721 AACTGGCTGTGCGCAACACTGGAATTTAGTGTTCAGAGCTTCAAGAGATCAATGCT 780
Qy 829 TTCAACTGGGCCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGCTTCTCAACAAGCTC 888
Db 781 TTTAATCTGGGCCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGCTTCTTATCAAGCTC 840
Qy 889 CGGCGAGATGCCAATATGTGCGTACAGATACCTTGAAATTCCTGATGCGCAGTGGGAT 948
Db 841 AACAACCTGAGTGAATGCTCATACAGGACACCTTGAGTTCCTGATGCGCGCCAGAC 900
Qy 949 TTCTGCAATCTCTTGGAAAAATCTGTGTGAACATCATGCGCTTCTTGAATCTTTTGA 1008
Db 901 TTCTGCAATCTTCTTGGAAAAATATGTGAGACCAATGCTTCTTCAAGACTCTTCGAA 960
Qy 1009 GAGCCCAACCAAAAGCCCAAGCCGTCCTCTTTAGCCGGGGTCAATCTTTCGGTTCACT 1068
Db 961 GAACCTAAGCCAAAGCCAAAGCCTGTTCTTTCAGCCGAGGGTCGTCTTCGGTTTACT 1020
Qy 1069 GGTCCGACTCAGAGAGAGTTCGACTATGTTAAAGAGAGACATAAGAGGTGAG 1128
Db 1021 GGTCCGACTCAGAGAGAGTTCGACTATGTTAAAGAGAGCGGCCCAAGAGAGGTGAG 1080
Qy 1129 TTTGAAAGAGACAGCAAGATTCATTCTATCCGAGCCTTGCTTCAAGCCTACAGA 1188
Db 1081 TTTGAAAGAGACAGCAAGATTCATTCTATCCGAGCCTTGCTTCAAGCCTACAGA 1140
Qy 1189 CTGAATTCGAAAGTGTGAGCAGTCTCAGCAGAGCAGCAGCTTCAATTTGGAGAGGT 1248
Db 1141 CGGAATTCAGAAAGTGCCTCAAAAAGTCTCCACAGAGCGCAAGCCTTACGTTGAGAGAGT 1200
Qy 1249 GCGCAATCTTCAGAGGGGCCAGAGCTCCGGCCGAGAGAAAGAACCAAGGTTCGCCGGG 1308
Db 1201 ACCGAGTCTTCGGGTGCGCAGAGCTGCGCAGAGAGAAAGCGGAGGCTGTACTTGG 1260
Qy 1309 GAGCCGGGGTGCACCCGAGCCCTGCGCGAGAGAAAGCGCGGGGTAAACAAGAGGCG 1368
Db 1261 GAGTTGGACCTTCACAAAGCCCGCCCTCCCTAAGAGCCTTCAGAGGAGCAAGCTGCA 1320
Qy 1369 GACGAGCCGCTCGGCGCCACGAGAGAA-----GAGAGAGAGGTCTTAAAGATAG 1422
Db 1321 GATGGAACCAACCGTGTGCGCGCTGAGAGAAAGAGAGAGAGGTGGCAAGAGCGGC 1380
Qy 1423 ACCCAGCAGATTAACCTCAGCCCCCGCAGCCCAAGCAGCTCCCTGATGCGCAGTCT 1482
Db 1381 ATCCCGCCAAAGTAACCTCAGCCCCCGCAGCCCAAGCAGAGCTCCCTGACTGTAGCCCT 1440
Qy 1483 CACCTTTCGAGCTGTCTGTGAACCTGCAAGGGGGAGGTGGCCCTGCAACGTAACCTTG 1542
Db 1441 CACCTTTCAGAGCTGTCTCACTCACTCAGGGAGAGAGCGCCCCAGCCCAACCTGACCTG 1500
Qy 1543 TCTCCCAACCTGAGCCCGACACCAAGAGGCTCTCCCTGTATCAAGCCGCTGCTGAAT 1602
Db 1501 TCTCCCAACCTGAGCTCTGACCAAGAGAGGCTCTCCCTGTATGAGCCGCTCTGAAC 1560
Qy 1603 GACCAAGGCTGCCCCCGAGCGACGATGAGATGAGAGGCCGAGAGAAAGATTTCCAACT 1662
Db 1561 GACCAAGGCTGCCCCCGAGCGACGACGAGAGAGAGAGAGAGAGAGATTTCCAACT 1620
Qy 1663 GATTAAGCGTACTTCTAGCTAAGAGAGTCTTACCAACGAGAGAACTATCTGAAGAT 1722
Db 1621 GACAAAGGCTACTACTTCTTAAGAGAGTCTTCAACCAAGAGCGACGTACTGAAGAGC 1680
Qy 1723 CTGGAAGTATCACTTCTGTGTTTCAAGACAAGTGAAGAAAGAGACGCAATGCCGAA 1782
Db 1681 CTGGAAGTATCTCTCTGTGTTTCAAGACAAGTGAAGAGAGTCTGAAGAGAGC 1740
Qy 1783 GCACTGAAAAGTCTCTATTTCCGAAATTTTGAACCTTTGCAAAATTTCTACTAATTTT 1842
Db 1741 GCCTTGAAAGTCTCAATATTCCTCAAAATTTTGAACCTTTGCAAAAGTTCAATCAATTTT 1800

Qy 1843 CTCAAGAAAATTTAGCAACGACTTGGCTGTGGGAAGGCGGCTCAATATGCCCAATCA-- 1900
Db 1801 CTGAAGAAAATTTAGCAACGACTTGGCTGTGGGAAGGCGGCTCAATATGCCCAATCA 1860
Qy 1901 -GAGATTAACAAAAGATGCGGATGTCTATGCTGAAGAACATTTAGGGGATGAAGACCTG 1959
Db 1861 GGAGATTAACAAAAGATGCGGATGTCTATGCTGAAGAAAATTTAGGGGATGAAGACCTG 1920
Qy 1960 GCGGCTCACTGTGGAACACAGAGGCGCTTGGAGGCGCTGGAAGATGAATCAAGAGC 2019
Db 1921 GCGGCGCACTGTGGAACACAGAGGCGCTTGGAGGCGCTGGAAGATGAATCAAGAGC 1980
Qy 2020 TCCCGCGGCTGGAAGAACCTTGTGACAGACTTTGAGCTGCAAGAGTGTGTTACTTACG 2079
Db 1981 TCGCGCGGCTGGAAGAACCTTGTGCGAGACTTTGAGCTGCAAGAGTGTGTTACTTACG 2040
Qy 2080 CTCAACACTTCTCTCTGCGGCACTGCAACCGGCTCATGCACTACAGCAGTCTTGAG 2139
Db 2041 CTCAACACTTCTCTCTGCGGCACTGCAACCGGCTCATGCACTATACATGCTCTGAG 2100
Qy 2140 CGGCTGTGCAACACACACCGGCGAGCGGCACTTCAAGGATGCGCAGCGGCTTGG 2199
Db 2101 AGGCTGTGCAACACACACCGGCGAGCTTCAAGGATGCGCAGAGCTGCGCTG 2160
Qy 2200 GCAAGATCAAGAGATGTGAGCAAGCTCAACGGTACGATGATCAAGATGAAGATTTT 2259
Db 2161 GCGAGATCAAGAGATGTGAGCAAGCTGATGATGATGATCAAGATGAAGATTTT 2220
Qy 2260 CAGAACTGTGCAAGAACTCAAGAAAGATTTGATGGAATTTGACATCTTGTGTTCCGGA 2319
Db 2221 CAGAACTGTGCAAGAACTCAAGAAAGATTTGATGGAATTTGACATCTTGTGATTCAGAG 2280
Qy 2320 AGGAGTTCATCCGCTGAGGAGGCTCTGAGCAAGCTCTGCGGGAAGGGGCTCCAGAGGCG 2379
Db 2281 AGGAGTTCATCCGCTGAGGAGGCTCTGAGCAAGCTCTGCGGGAAGGGGCTTCAGAGGCG 2340
Qy 2380 ATGTCTTCTCTGTCAAGAGAGTCTCTGTATACAGAGCCGAGGCTGACGCGCTTCAAT 2439
Db 2341 ATGTCTTCTCTGTCAAGAGAGTCTCTGTATACAGAGCCGAGGCTGACGCGCTTCAAT 2400
Qy 2440 CAGTTAAAGTTCACGGGAGCTCCGCTCTATGAGCAGTGAAGATTTGAGAGAGCGAAGC 2499
Db 2401 CAGTTAAAGTTCACGGGAGCTCCGCTCTATGAGCAGTGAAGATTTGAGAGAGCGAAGC 2460
Qy 2500 GAGTGGGGGGTCCCACTGACTGACCTCGGGGCGAGCGGAGTGCATCATGCTGAGCC 2559
Db 2461 GAGTGGGGGGTCCCACTGACTGACCTCGGGGCGAGCGGAGTGCATCATGCTGAGCC 2520
Qy 2560 GCGAGTTCCTGAGTCCGAGATGAGAGATGAGTTGAGACATTCAGAGTGCATTTGACCTG 2619
Db 2521 GCGAGTTCCTGAGTCCGAGATGAGAGATGAGTTGAGACATTCAGAGTGCATTTGACCTG 2580
Qy 2620 GCGGAGAGAGCAGAGAGCCCGGCTCTGAGTTCTGAGCAGAGGCGCCCTGCAACAAG 2679
Db 2581 GCGGAGAGAGCAGAGAGCCCGGCTCTGAGTTCTGAGCAGAGGCGCCCTGCAACAAG 2640
Qy 2680 TCCCGTGAAGGCCACCGGGCTGACAGAGATGAGAGATGACCTGAGGCGCTCGGCG 2739
Db 2641 TCCCGTGAAGGCCACCGGGCTGACAGAGATGAGAGATGACCTGAGGCGCTCGGCG 2700
Qy 2740 ACATGCTGAGAGCGCCAGAGCCCGGCAACCAATGTGTCAAGTGTGCTGAGC 2799
Db 2701 ACATGCTGAGAGCGCCAGAGCCCGGCAACCAATGTGTCAAGTGTGCTGAGC 2760
Qy 2800 GCGAACAACAGGCTTCCATGATGAGACTTCAAGCTGCAAGTGAAGATCAATGTTCTGGA 2859
Db 2761 GCGAACAACAGGCTTCCATGATGAGACTTCAAGCTGCAAGTGAAGATCAATGTTCTGGA 2820
Qy 2860 AACCTGTGAAGAAATTCAAAACAGCAACGGGTGCGAAGCTGTGGGTGTGTTCA 2919
Db 2821 AACCTGTGAAGAAATTCAAAACAGCAACGGGTGCGAAGCTGTGGGTGTGTTCA 2880

QY 2920 AACTTTCCTGCTTCTTCTTCTAATCATCACACAGAGCAATCATCCCTTGGCAGCTCCCT 2979
 Db 2881 AACTTTCCTGCTTCTTCTTCTAATCATCACACAGAGCAATCATCCCTTGGCAGCTCCCT 2940
 QY 2980 CTGCTGGGCTACTGCTGCTCACCATTCCCTCTGAGTCCGAAACATCCAGAAAGACTACGTG 3039
 Db 2941 CTCTGGGCTACTGCTGCTCACCATTCCCTCTGAGTCCGAAACATCCAGAAAGACTACGTG 3000
 QY 3040 TTCAAGCTGACTTCAAGTCCCAAGCTCTACTACTCTGAGGCGGAAAGCGAGTACAGTTTC 3099
 Db 3001 TTCAAGCTGACTTCAAGTCCCAAGCTCTACTACTCTGAGGCGGAAAGCGAGTACAGTTTC 3060
 QY 3100 GAAAGGTGATGAGATGATCCGACAGTCCAGCTCTGCTCGACCCCACTGTTTG 3159
 Db 3061 GAAAGGTGATGAGATGATCCGACAGTCCAGCTCTGCTCGACCCCACTGTTTG 3120
 QY 3160 AGCCACAAAGAGTCTCTTGTGTATGA 3186
 Db 3121 AGTCACAAAGATCTCAGTCTACTGA 3147

RESULT 5
 AK171713
 LOCUS
 DEFINITION
 AK171713 3743 bp mRNA linear HTC 21-SEP-2005
 Mus musculus activated spleen cDNA, RIKEN full-length enriched library, clone:FB30006D05 product:FERM, Rhogef and pleckstrin domain protein 2, full insert sequence.
 AK171713
 VERSION
 AK171713.1 GI:74152707
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 2 10349636
 3

JOURNAL
 PUBMED
 11042159
 TITLE
 AUTHORS
 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Komano, H., Akiyama, J., Nishii, K., Katsunari, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishihara, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 4 11076861

JOURNAL
 PUBMED
 11076861
 TITLE
 AUTHORS
 Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukunishi, Y., Komano, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishii, K., Kiyosawa, H., Kondo, S., Yamamoto, R., Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Cavaletto, T., Saito, R., Kato, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kato, T., Lewis, S., Matsuo, Y., Nishii, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nishii, C., King, B., Kochiwa, H., Quackenbush, J., Schriml, L.M., Straub, F., Suzuki, R., Tomita, M., Wagner, L., Maehiro, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barch, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,

Marchionni, L., Maehiro, J., Mazzarelli, J., Momberts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S., and Hayashizaki, Y.
 RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 11217851

CONSRMT
 TITLE
 JOURNAL
 PUBMED
 11217851

REFERENCE
 AUTHORS
 5
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishii, C., Saito, R., Saito, R., Suzuki, H., Yamamoto, R., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, I., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanari, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusa, V., Choctha, C., Corbani, L.E., Cousins, S., Dalla, E., Dragan, T.A., Fletcher, C.F., Forrest, A., Fraser, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvie, E.D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, M.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayashizaki, N., Hirokawa, N., Kishikawa, T., Komano, H., Nakamura, M., Sakazume, N., Saito, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Aizawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E., and Hayashizaki, Y.
 FANTOM Consortium
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420 (6915), 563-573 (2002)
 12466851

CONSRMT
 TITLE
 JOURNAL
 PUBMED
 12466851

REFERENCE
 AUTHORS
 6
 Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Fith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aldins, V., Allen, J.E., Ambesi-Imbondato, A., Apweiler, R., Attaluri, R.N., Bailey, T.L., Bansal, I., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Chiswick, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashizaki, Y., Hensch, T.K., Hirokawa, N., Hill, D., Hummel, K.L., Iacono, M., Ikeno, K., Iwano, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsunaga, S., Miki, H., Mignone, F., Miyake, S., Morita, K., Mottagui-Faber, S., Mulder, N., Nakano, N., Nakaguchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rosch, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekizawa, K., Semple, C.A., Sene, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Sill, D., Sinclair, B., Sperling, S., Stupka, E., Sugizaki, K., Sultana, R., Takenaka, Y., Takai, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R.,

CONSRM	FANTOM Consortium
TITLE	The transcriptional landscape of the mammalian genome
JOURNAL	Science 309 (5740), 1558-1563 (2005)
PUBMED	16141072
REFERENCE	7
AUTHORS	Katayama, S., Tomaru, Y., Kaeukawa, T., Waki, K., Nakanishi, M.,

AUTHORS
 Katsuyama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakamichi, M.,
 Nakamura, M., Nishida, H., Yap, C. C., Suzuki, M., Kawai, J., Suzuki, H.,
 Cerniglia, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T.,
 Peng, K. C., Hallinan, J., Mattick, J., Hume, D. A., Lipovich, L.,
 Batalov, S., Engstrom, P. G., Mizuno, Y., Feghali, M. A., Sanderlin, A.,
 Chalk, A. M., Mottagui-Tabar, S., Liang, Z., Lemhard, B. and
 Mhalestedt, C.

CONSRM RIKEN Genome Exploration Research Group
TITLE Antisense transcription in the mammalian transcriptome
JOURNAL Science 309 (5740), 1564-1566 (2005)

AUTHORS
 Arkhava, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K.,
 Hori, F., Iida, Y., Imamura, K., Imochani, K., Itoh, M., Kanagawa, S.,
 Kawai, J., Kojima, M., Kono, H., Murata, M., Nakamura, M., Nishitani, A.,
 Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D.,
 Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Wataniki, A.,
 Muramatsu, M., and Hayashizaki, Y.

Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Saito-cho, Tsukuba, Ibaraki, 305-8565, Kanagawa, 230-0045, Japan (e-mail: genome-research@gscc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel.81-45-503-9232, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genetic Sciences Center and Genome Science Laboratory in Riken, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for full details.
URL: <http://genome.gsc.riken.jp/fantom>
URL: <http://fantom.gsc.riken.jp/>

FEATURES	Location/Qualifiers
source	1. .3743

1030 CCCGCTCTTTAGCCGCGGGGATCATTTCCGTTTCAGTGTGCGACTCAGAAAGAGTT 1089
1129 GCTGTCTTCTTACGCGGAGCTCTCTTCAGATACGTGAAAGAACTCAGAAACAACTA 1188
1090 CTCGACTGTTTAAAGAGAGACATTAAGAGGTGAGTTTGAAGAAAGACACAGCAG 1149
1189 GTAGATTATGTCAAAAGACGCTGGAAATGAAGAAATTCATATCAGAAAGACGCGACATG 1248
1150 ATTCTCTTCCGAGACCTTGTCTTCAACGCTTACAGATCTGAATTTGGAAAGTCTG 1206
1249 ACTCGACATCTCTTCACTGCTCTGATCTGATGATCTGCTTAAACAGACGCTCTCTTACC 1308
1207 GAGCAGCTCTCAGAGAGACACAGCCTTACATTTGAGAAAGTCCGCAATCTCCAGGGGCG 1266
1309 GATGGCTTGAAGAACTTCTGCTCTCTCTCTTCAAGCAAAATGCTCTCTTTATCTC----- 1361
1267 CAGAGCTGCGCGGAGAAAGAAACCGAAGTTTCCGCGGAGAGCCGCGGCTGCACTCCG 1326
1362 ---ACCCCTAGTTCTTCTTGTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1418
1327 AGCCCTGCGCGGAGAGAAAGCCCGGGGTAAACAAGAGCGGACGAGCCGCTCGGCG 1386
1419 CTCCTTGTGATCCGCGAGCTCCGCTCATTAAGAGACAGACAGAGAGAGAGAGAGCGG 1478
1387 CCCACGAGAGAGAGAGAGAGTCTTAAAGATAGAACCCAGCAGAGTAAACCTCAGCCC 1446
1479 ACCA--TCACTACTGACGCGCCGACAGACAGTCCGCTCATCTCTGAGACCCCGCTG 1536
1447 CCGCAGCCAAAGACAGGCTCTCTGATGCGAGTCTCTCACTTTCCGAGCTGTCTGTAAAC 1506
1537 CTCGCGCTGTGTCAAGGCTTTTCTAATGATGCTCTCAGCCTTCTCCCTCAGCCTGAAAG 1596
1507 TCGCAGGGGGAGTGGCCCTCTGCAACGTGACCTTGTCTCCCACTGAGCCCGGACCC 1566
1597 AGCCACCTGAGCTGTGCTCTGAGCTTCAAGCCGCG-----CTGATTAACGCTGAG 1647
1567 AAGCAGAGCTCTCTCTTGTATCAGCCGCTGCTGAATGACCAAGCCGCTGCGCGAGAGAC 1626
1648 CAGGGTGCATCTCTGTGTCTCAAGCTCTGTGTCTCAAGTGGCCCTGATGAGCAGAGTGGAC 1707
1627 GATGAGATGAGGGCGGAGAGAGATTCCTCACTGATTAAGCGTATCATAGCTAAG 1686
1708 AACCAAGAGAGAGAGAGACACAGACATGCGAAGATGAGGCTTATTCATAGCAAG 1767
1687 GAAAGTCTTACCAACCGAGAGAACTATCTGAAGATCTCGAATTTCACTTGTGTGTTT 1746
1768 GAGATTCTGCTTACAGAACGAACTATCTGAAGATTTAAGATTATCACTGTGTGTTT 1827
1747 CAGAGCAGTGAAGAGAGAGAGAGAGATGCGGAGACCTGAAGAGCTGAAGAGCTCATATTC 1806
1828 CGAGCGCTGTATTAAG 1887
1807 AATTTTGAACCTTTGACAAATTTCACTAATTTTCTCAAGAGAAATGAGCAACACTT 1866
1888 AACATTGATCTCACTTACAGATTCACAGAGGCTCTCTTCAAGAGGTGAAGAGAGAGCTG 1947
1867 GCGCTGTGAGAGAGCGCTCAATTCGCCA---AATCAGAGATTTACCAAGAGATCGCGAT 1923
1948 GCACTTGTGAG 2007
1924 GTCATGCTGAAGAACTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1983
2008 ATCTCTCTCAG 2067
1984 GAGGCTTTGAG 2043
2068 GAGGCTTAAACAAAG 2127
2044 AAGAGACTTTGAG 2103
2128 AAGAGATTTGAG 2187
2104 CTGACAGGAG 2163

2188 GTCCAGAGGCTATCCACTACCGTCTGTGCTGAGCGGCTGTGTCTCACTACTCTCCCT 2247
2164 AGCCACGCGGACTTCAAGGAGCTGCGGAGCCGCTTTGACAGAGATCAACGAGATGTGGCA 2223
2248 GGGCACCGGAGCTATGCGGAG 2307
2224 CAGCTCAAGGTATGATGATCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2283
2308 GAGCTCAG 2367
2284 GATTGATTTGAG 2343
2368 GACCTGTGTGTATGAG 2427
2344 CTCAGAGAGCTCTCGGAG 2403
2428 CTGCAAGAGCTCAAG 2487
2404 CTGCTATACAG 2463
2488 TTACTGTATTAAG 2547
2464 CCGCTTATGAG 2523
2548 CACCTCGGTGAG 2607
2524 ACCCTCGGAG 2583
2608 ACCATTAATGAG 2667
2584 AAGTGGTTGAG 2643
2668 AAGTGGATGAG 2727
2644 CTTGAGTCTGTGCGCAG 2703
2728 CAGTGTCTGTG---GAGGCGCGGTGTATCTCGTACCTCTGATCTTCTGATGAGATC 2784
2704 GACAGAGAGTCAAGAGATGATCTGAGCGCTGCGGACATCTGCTGAGAGCGGAGCGCGG 2763
2785 TCTCTGAG 2844
2764 CACCGGAG 2823
2845 CACCGGAG 2904
2824 GACTTCAAGATGAG 2883
2905 GACCAAG 2964
2884 AGCAAG 2943
2965 AGTAAATGCTGAG 3024
2944 TCAACCAAG 3003
3025 ACACATCAG 3084
3004 CCGTGTAGTCCGAG 3063
3085 CCGAG 3144
3145 GTCTACTTCTTCCGCGGTGAG 3204
3124 AGTGCACAG 3183
3205 AAGGAG 3264
3184 TGATGCGCGAG 3223

Db 3265 CCAGGCTGAGGACAGATCAGAGAAAGAGCATGCC 3304

RESULT 6
AK050860 3874 bp mRNA linear HTC 02-SEP-2005
Mus musculus 9 days embryo whole body cDNA, RIKEN full-length
enriched library, clone: D030026M03 product: weakly similar to CDEP
[Homo sapiens], full insert sequence.

ACCESSION
AK050860
VERSION
AK050860.1 GI:26094164
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
PUBMED
10349636
REFERENCE
AUTHORS
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
PUBMED
11042159
REFERENCE
AUTHORS
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikigami, T., Kashiwagi, K.,
Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexed capillary sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
PUBMED
11076861
REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409, 685-690 (2001)
PUBMED
11542000
REFERENCE
AUTHORS
5
The FANTOM Consortium, the RIKEN Genome Exploration Research Group
Phase I and II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)
PUBMED
12000000
REFERENCE
AUTHORS
6
RIKEN Genome Exploration Research Group, Genome Science Group
(Antisense Network Core Team) and the FANTOM Consortium.
TITLE
Antisense transcription in the Mammalian Transcriptome
JOURNAL
Science 309, 1564-1566 (2005)
PUBMED
15840000
REFERENCE
AUTHORS
7
The FANTOM Consortium, Riken Genome Exploration Research Group and
Genome Science Group (Genome Network Project Core Group).
TITLE
The Transcriptional Landscape of the Mammalian Genome
JOURNAL
Science 309, 1559-1563 (2005)
PUBMED
16000000
REFERENCE
AUTHORS
8 (bases 1 to 3874)
Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P.,
Hayashida, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hukushima, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsumi, N.,
Okazaki, Y., Saito, K., Saitoh, H., Sakai, C., Sakai, K., Sakamoto, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL
Direct Submission
Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa, 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome-gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome-gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
URL: http://fantom.qualifiers

FEATURES
source
1..3874
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:DB:D030026M03"
/db_xref="taxon:10090"
/clone="D030026M03"
/tissue_type="whole body"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="9 days embryo"
1..3874
/note="putative
weakly similar to CDEP (Homo sapiens) (SEPR|09Y4F1,
evidence: FASV, 53.7%id, 99.7%length, match=3125)"

misc_feature
Query Match 29.4%; Score 1013.6; DB 6; Length 3874;
Best Local Similarity 59.0%; Pred. No. 4.3e-260;
Matches 1864; Conservative 0; Mismatches 1269; Indels 29; Gaps 6;

ORIGIN
70 CCGAGCCGAGATCAACGCTGGGGCCCGGAAAATTCGGGATGATGATCTGGAAGCT 129
|||
103 CCACCTTACAGGACACGCTTGGTGCCAGACCCGATGAGTACCTTGAGGCA 162
|||
130 GGACAGAGCCGCGCCCAACACTTCAGAAACCTGTCATCAATCCAGATGCTG 189
|||
163 GAGCAGAGCTGTACCCAGATGAGAGAGACATGCCATGAGTGAAGTCTGCTG 222
|||
190 GATGACCCAGAGGACATTTGAAGTCCACAAAGAGCTCTGGAGAGTCTGAT 249
|||
223 GACAGCAGTGAAGTATTTGACATCGAGCCGAAGTGAAGGAGGAGTCTTACTGACT 282
|||
250 GCAATTCACACCACTCACTGCTGGAAGTGAAGTATTTGGCTGAGATTTCTGAT 309
|||
283 CAAGTGAAGACATTTAATCTCAATGAATGTACTTGTGCTGAGATTCAGAAAT 342
|||
310 CACAAAAGATCAACGCTGCTGATCTCTTAAACCATTTGTGAACAGATTTGAAG 369
|||
343 GTCACGCTCTACTGATTTGGCTTGAACCTATGAACCATCATATAGGCAAGTGAAG 402
|||
370 CCAAGACAGTTGTTGTTAAGTTGTGTAATTTCTTCCGCTGACACACACAATC 429
|||
403 CCAAAAATGCGGTCTTCGCTGCAATTAATTTTCCGCGCTGACCTGTCAGCTG 462
|||
430 CAAGAGAACATCAAGATTAATCTTGGCGCTGCAAGGAGAGAGAGCTGCTCAAGGC 489
|||
463 CAAGAGAGTACACAAAGACCTGTTTCCCTTGCACTCAAGAGAGAGAGCTTGAAGAA 522
|||
490 AGGTGAAGTGAATGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 549
|||
523 CGCCTGACCTGACAG 582
|||
550 ATTGGGAGATTTGATGAAGCTTGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 609
|||

Db 583 ATGCGAATTATGTAAGAAACCTTGATCGAAGAACCTCAAGCAATGATATTGGCC 642
Qy 610 CACGAGAACCACTAGAGGACAAAAATGTTGGAATTTTCAACAATACCATTTGACAAACA 669
Db 643 AACGAGAGAAATCTTAGAAAAGATCTAGACTTTCATCAGAGGACACGGGCGACACT 702
Qy 670 CCAGCAAAATCAATTTCCAGCTCCTTAGAATTTGCCGTGAGATGATATGAAATC 729
Db 703 CCGGCAAGTCAGATTTCCAGGTCCTTAGATTTGCAAGAAAGCTGGAAATGATATGCAATC 762
Qy 730 CGGTTCACCCGGCCAGAGCAGGAAAGGACGAAATCAATCTGCGCTTGCACACG 789
Db 763 AGATTTCATGAGCTTTCAGAGAGAGGACCAAGATTAATCTAGAGATTTCTCACATG 822
Qy 790 GGAATTTCAATGTTTTCAGGGTTTCACTAAGATCAATGCTTCACTGGGCGAAGTCCGG 849
Db 823 GGTGTCTGGTGTTCAGGGTACCAACAAATCAATCTTCAACTGGGTCAAAAGTCCGT 882
Qy 850 AAGCTAGCTTCAAGAGAGAGCGCTTCTCATCAAGCTCCGGCCAGATGCAATAGTGC 909
Db 883 AACTAAGCTTCAAGAGAAAGATTTCTTATCAAACTCCACCTGAGTCCATGGGCC 942
Qy 910 TACCAAGATACCTTGAATTCCTGATGCGAGTCGGGATTTCTGAAGTCTTCTGAAA 969
Db 943 TACCAAGACACGTTAGAGTTCTTGTGGGTAGCAGAGATGAATGTAAAGACTTCTGGAAG 1002
Qy 970 ATCTGTGTTGAACATCATGCTTCTTTAGCTTTTGAAGGCCCAACCAAGCCCAAG 1029
Db 1003 ATATGTGTGAGTACCACTACCTTTTGAAGCTCTGAGCAGCTTAAAGCCAAAGGCAAG 1062
Qy 1030 CCGGTCTCTTAAAGCGGGGTATCATATTCCGTTCAAGTGTGAGCTCAGAGCAGGTT 1089
Db 1063 GCTGTCTTCTTCAAGCCGAGGCTCTCTTCAATACGATGAGAAAGCTCAGAAACAATCA 1122
Qy 1090 CTCGACTATGTTAAAGAGGAGACATTAAGAGTGTGATTTGAAGAAAGACACAGCAAG 1149
Db 1123 GTAGATTATGTCAAGACGGTGGAAATGAAGAAATTCATACGAAAGACGGCACAGTAAG 1182
Qy 1150 ATTCT--ATTCTATCCGAGCTTGTCTTCAAGCTTACAGAACTGAATTTGGAAATGCTG 1206
Db 1183 ACTCGGACATCTTCAATGCTCTGATCTGTAGATCTGCTTAAACAGAGCGTCTCTTCAAC 1242
Qy 1207 GAGCAGCTCAGAGAGACACGAGCTTACATTTGAGAAAGTGCAGAACTCCAGGGGGC 1266
Db 1243 GATGGCTTGAAGACTTCTGCTCTCTCTCTCTTCAAGCAAAATGCTCTTTATTC----- 1295
Qy 1267 CAGAGCTGCGGCGAGGAAAGAACCGAAGTTTCCGCCGAGGAGCGGGGTGCAACCG 1326
Db 1296 ---ACCCCTAGTTCTTCTTGTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1352
Qy 1327 AGCCCTGCGCGAGAGAGAGCCCGCGGGTTAACAAGAGCGGACCGAGCCGCTGGCG 1386
Db 1353 CTCCTTGTGATCCCGAGGCTCCCGTCAACAAGAGACGACAGCAGAGAGGAGCAGCGG 1412
Qy 1387 CCCACGAGGAGAGAGGAGGTGTTAAGATAGAGCCGAGAGAGTAACCTCAGCCC 1446
Db 1413 ACCA--TCATCATCTAGCGGCCACAGACACAGTGGCCCATCTCCCTGAGCCCCCGTG 1470
Qy 1447 CCGAGCCAAAGACAGGCTCCCTGATCTGAGCTTCAACCTTTCCAGAGCTGTCTGTAAC 1506
Db 1471 CTCGGGCTGTGTCAGAGCTTTTCTATGGAATGCTCTAGGCTTCTCCCTCAGCCTGAG 1530
Qy 1507 TCGCAGGGGAGAGTGGCCCTTGCACAGTGACTTGTCTCCCAACTGAGCCCGACACC 1566
Db 1531 AGCCACTGAGCTGTGCTCTGAGCTTCAAGCCGCG-----CTGAATACAGCTGAG 1581
Qy 1567 AAGCAGGCTCTCTCTTATCAGCCGCTGTGAATGACAGGCTGAGCCCGCCGACGAGC 1626
Db 1582 CAGGGTGCATCCCTGTGTCTCAGCCCTGTGTCTCAGTGGCGCTGTACAGCAGAGATGAC 1641
Qy 1627 GATGAGGATGAGGCGCGAGGAAAGATTTCCAACTGATTAAGCTTCAATAGCTAAG 1686
Db 1642 AACCAAGAGAGAGAGAAACAAAGACATGCCAGAAAGATGAGGCTTATTTCAATAGCCAAG 1701

Qy 1687 GAATGTTTACCAACGAGGGAACATATCTGAAGGATCTCGAAGTTATCATTTGTTGTTT 1746
Db 1702 GAGATTTCTGCTTACAGAAAGAACCTTATCTGAAGGATTTTGAAGTTATCATCTGTGTGTT 1761
Qy 1747 CAGAGCAGATGAGCAAGAGAGAGCGCATGCGGAAAGACCTGAAAGTCTCATATTTCCG 1806
Db 1762 CGGAGGCTGTATCAAGGAGGAGGCGATGCTCGAGCCCTGATGAGCCCTGCTTTCTTC 1821
Qy 1807 AATTTGAACCTTTGCACAAATTTCAATCTAATTTTCTCAAGGAAATTTGACACACAT 1866
Db 1822 AACATTAATCTAGTCTACAGATTCACAGAGGCTTCTTCAACAGAGTGAACAGAGCTG 1881
Qy 1867 GCGCTGGGAAAGGCGCTCAATGCGCA--AATCAGAGATTACCAAGAAATCGCGAT 1923
Db 1882 GACCTCGGAGAGGCGCTTCTGCTGCTCACTTAAAGTATCAGAGGATTCGAGATTCGGGAC 1941
Qy 1924 GTCATCTGAGAGAACATTCAGGGCATAGAGCACTTGCGGCTCACCTGTGAGAGCACAG 1983
Db 1942 ATCTCTCAGAGAACTGCGTCACTTAAAGGAAATTTACTAGCTATCCAAAGACAGAT 2001
Qy 1984 GAGGCTTGAAGCCCTGAGAAATGGAATCAGAGCTCCCGGCGCTGAGAACTTCTGC 2043
Db 2002 GAGTCTTAACAGAACTGGAAGAAAGCCACAAACACTGTAAAGCTGAGGCGAGCTTAC 2061
Qy 2044 AAGACCTTGAAGTGAAGAGTGTGTACTACGCTCAACACCTTCCCTGCGGCGCA 2103
Db 2062 AAGAGTTTGAAGTCCAAAGGCTGTCTACCTGCTCTCAACATTTCTGCTGAAGCCC 2121
Qy 2104 CTCACCGGCTCATGACTTACAGAGCTCTGAGCGGCTGTGCAACACACCGCG 2163
Db 2122 GTCCAGAGGCTGTACACGTGTGCTGTGAGCGGCTGTGTCTCATCTCTCT 2181
Qy 2164 AGCCACGCCGACTTACAGGAGCTGCCAGCGCTTTGCGAGATTCAGAGATGTGGCA 2223
Db 2182 GGGCACCGGAGCTATGCCACCTGCACTGCACTGAGGCACTGAAAGCCATCAAGAGTACAC 2241
Qy 2224 CAGCTCAGGTTACATGATCAAGATGGAAGATTTCCAGAGCTGACAGACTCAGAAA 2283
Db 2242 GAGCTCAGCAAGACCTTACCGGCTGGAAGACCTACAGAAATTCAGAGGCTCAGACA 2301
Qy 2284 GATTTGATTTGCAATGACATCTTGTGTTCGCGAAGGAGTTCATCCGTCTGGCAGC 2343
Db 2302 GACCTGTGTGTATGAAGAACTCATTTGCTCTGGGAGGAGATTTATCCGTGAGGCTGC 2361
Qy 2344 CTCAGAGCTCTCGGGAAAGGGGCTCAGAGCGCATGTTCTTCTGTTCAACGACGTC 2403
Db 2362 CTCACAGACTCACAAAGAGGGGCTGCAACAGAGATGTTTTTCTGTTCTCAGATATG 2421
Qy 2404 CTGCTATACAGAGCGGGGGCTGACGGCTCCATCAATCAATTAAGTCCAGCGGCGCTC 2463
Db 2422 TTGCTGTATACAGAAAGTGTACAGAGGCGAGTCAATTCGAGTCCGTGCTTCTT 2481
Qy 2464 CCGCTTATGCAATGCAATTTGAGAGAGCAGAGCAGATGGGGGTGCCCCATGCTG 2523
Db 2482 CCACTCCGTGCAATGTGTATGAAGAAATGAGAAATGATGTGTCTCTCATTTGCTTC 2541
Qy 2524 ACCCTCGGGCCAGCGGCAATCATCTGTCGCGCAATTTCTGTGTCGAGATGAG 2583
Db 2542 ACCATTAATGAGCTCAGAAAACAATTTGTGAGCAGCAGACCTCGCTAAGAAAGGAA 2601
Qy 2584 AAGTGTGTAAGAGACATCAAGATGCGCATTTGACCTGTGGGAGAGAGCAGAGCCCGG 2643
Db 2602 AAGTGAATCAGAGACTTAATGCAAGAAATTCAGAGCAGCAAGCTATCGGTACTACCC 2661
Qy 2644 CCTGAGTTCTGCGCAGAGCCCGCTGCAACAAGTCCCTGATGAAGCCACCGGCGCT 2703
Db 2662 CAGGTGTGTGAGAGCGCGGTATTACTGTACCCATGAATCTTGTATGAAGTCTCT 2721
Qy 2704 GACAGAGTCAAGAGATGACTGAGCGCTTGCAGACATGCTGAGCGCAGCGCCCG 2763
Db 2722 CTGGAAGAAATCAAGATGTTGAGGAGAAAGCGGGGCTGCTGAGGGGAAACGCGCAG 2781

Db 960 TTACTTACC 968

|||||

RESULT 8
BX354921 1000 bp mRNA linear EST 23-APR-2004
LOCUS BX354921 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DC027YA10 5-PRIME, mRNA sequence.
ACCESSION BX354921 GI:46553760
VERSION BX354921.2
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1000)
Li,W.B., Gruber,C., Jessee,J. and Polayer,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30383921.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7229.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DC027BA050P1c=7229.f.
Location/Qualifiers
1. 1000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC027YA10"
/isue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source

ORIGIN

Query Match 27.3%; Score 940.2; DB 4; Length 1000;
Best Local Similarity 99.2%; Pred. No.1.5e-240;
Matches 961; Conservative 4; Mismatches 2; Indels 2; Gaps 2;

1110 AGGACATAAGAGAGTGCAGTTTGAAAGAGACACAGCAAGATTCATTTCGAGGCT 1169
1 AGGACATAAGAGAGTGCAGTTTGAAAGAGACACAGCAAGATTCATTTCGAGGCT 60
1170 TGGTTACAGCCTACAGAACTGAAATTCGGAAGTCTGAGCACTTCAGACAGCACAG 1229
61 TGGTTACAGCCTACAGAACTGAAATTCGGAAGTCTGAGCACTTCAGACAGCACAG 120
1230 CCTTACATTTGAGAGAGTGCAGATCTCCAGAGGGGCGAGAGTGCAGGAGAAAGGA 1289
121 CCTTACATTTGAGAGAGTGCAGATCTCCAGAGGGGCGAGAGTGCAGGAGAAAGGA 180
1290 ACCGAAAGTTTCGCGCGGGAGGCGGGGCTGCACCCGAGCCCTGCGCCGAGAGAAAGCC 1349
181 ACCGAA-GTTTCGCGCGGGAGGCGGGGCTGCACCCGAGCCCTGCGCCGAGAGAAAGCC 239
1350 CCGGGGTAAACAAGACGAGCGGAGCCGCTCGGCGCCCAAGAGAGAAAGAGAGGT 1409
240 CCGGGGTAAACAAGACGAGCGGAGCCGCTCGGCGCCCAAGAGAGAAAGAGAGGT 299
1410 CGTTAAGATAGAGACCAAGACAGTAACCTCAGCCCGCAGCAAGCAGAGCTCCCT 1469

Db 300 CGTTAAGATAGAGACCAAGACAGTAACCTCAGCCCGCAGCAAGAGAGGCTCCCT 359
1470 GACTGCACTCTCTACCTTCCGAGCTGTCTGTGAACCTGGCAGGGGAGATGGCCCTTC 1529
360 GACTGCACTCTCTACCTTCCGAGCTGTCTGTGAACCTGGCAGGGGAGATGGCCCTTC 419
1530 CAACGAGACTTGTCTCCCAAGCTGAGCCCGACACCAAGCAGGAGCTCCCTGTATCAG 1589
420 CAACGAGACTTGTCTCCCAAGCTGAGCCCGACACCAAGCAGGAGCTCCCTGTATCAG 479
1590 CCGGCTGTGAATGACACAGAGCTGCCCCCGGACGAGATGAGATGAGGCGCGAGAA 1649
480 CCGGCTGTGAATGACACAGAGCTGCCCCCGGACGAGATGAGATGAGGCGCGAGAA 539
1650 GAGATCCCAACTGATTAAGCTGATTCATTAAGTAAAGAGTGTCTTACACGAGGAA 1709
540 GAGATCCCAACTGATTAAGCTGATTCATTAAGTAAAGAGTGTCTTACACGAGGAA 599
1710 ATATCTGAAGAGATCTGAGATTAATCACTTCGTGTTTCAAGACACAGTAAAGAGA 1769
600 ATATCTGAAGAGATCTGAGATTAATCACTTCGTGTTTCAAGACACAGTAAAGAGA 659
1770 CCGCATGCCGAGACACTGAAAGTCTCATATTCGCAATTTTGAACCTTTCACAAAT 1829
660 CCGCATGCCGAGACACTGAAAGTCTCATATTCGCAATTTTGAACCTTTCACAAAT 719
1830 TCATCTAATTTTCTCAGAGAAATTGAGACAGACTTCCCTGTGGAGGCCGCTCAA 1889
720 TCATCTAATTTTCTCAGAGAAATTGAGACAGACTTCCCTGTGGAGGCCGCTCAA 779
1890 TGGCCAAATCAGAGATTAACAAAGATGGCGCATGCTGAGAGAAATTCAGGAGCAT 1949
780 TGGCCAAATCAGAGATTAACAAAGATGGCGCATGCTGAGAGAAATTCAGGAGCAT 839
1950 GAAGCACTGGGCGCTCACTGTGAGACAGACAGAGCC-TTGGAGGCCCTGGAGAAATG 2008
840 GAAGCACTGGGCGCTCACTGTGAGARCAAGCAGAGGCTTTGGAGGCCCTGGAGAAATG 899
2009 GAATCAAGAGCTCCGCGCGCTGGAGAACTTTCGACAGACTTTGAGCTGCAGAGGTGT 2068
900 GAATCAAGAGCTCCGCGCGCTGGAGAACTTTCGACAGACTTTTGTGCGAGAAAGTGT 959
2069 GTTACTAC 2077
960 GTTACTAC 968

RESULT 9
BX354920/c 1072 bp mRNA linear EST 23-APR-2004
LOCUS BX354920 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DC027YA10 3-PRIME, mRNA sequence.
ACCESSION BX354920
VERSION BX354920.2 GI:46552853
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1072)
Li,W.B., Gruber,C., Jessee,J. and Polayer,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30381927.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7229.f
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?c=CS0DC027BA05NP1&c=7229.f>.

FEATURES

source

```
1..1072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC027YA10"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match 27.1%; Score 933.4; DB 4; Length 1072;

Best Local Similarity 97.8%; Pred. No. 1.1e-238;

Matches 979; Conservative 4; Mismatches 14; Indels 4; Gaps 4;

```
2279 AGAAGATTGATTGCGATTGCAATCTTGTTCCGGGAAGGAGTTATCCGCTCG 2338
      |||
      |||
      |||
998 AARAAAGATTATTCATTBACAACTTGTTGTT- GGGGAAAGAGTTATCCGCTCG 940
      |||
      |||
      |||
2339 GCAGCTCAGCAAGCTCTCGGGGAAGGGCTCCAGAGGGCATGTTCTTCGTTCAAG 2398
      |||
      |||
      |||
939 GAGGCTCAGCAAGCTCTCGGGGAAG- GGGCTCCAGAGGGCATGTTCTTCGTTCAAG 881
      |||
      |||
      |||
2399 AGCTCTGCTATACAGAGCGGGGCTGACGGCTCCATCAGTTTAAAGTCCAGGGC 2458
      |||
      |||
      |||
880 AGCTCTGCTATACAGAGCGGGGCTGACGGCTCCATCAGTTTAAAGTCCAGGGC 821
      |||
      |||
      |||
2459 AGCTCTGCTATACAGAGCGGGGCTGACGGCTCCATCAGTTTAAAGTCCAGGGC 2518
      |||
      |||
      |||
820 AGCTCTGCTATACAGAGCGGGGCTGACGGCTCCATCAGTTTAAAGTCCAGGGC 762
      |||
      |||
      |||
2519 GCTGACCTCTCGGGGCGGCGGCGGCTCCATCAGTTTAAAGTCCAGGGC 2578
      |||
      |||
      |||
761 GCTGACCTCTCGGGGCGGCGGCGGCTCCATCAGTTTAAAGTCCAGGGC 702
      |||
      |||
      |||
2579 TGGAGAGTGGGTTGAGGACATCCAGATGGCCATTGACCTGGCGGAGAGAGCAGCC 2638
      |||
      |||
      |||
701 TGGAGAGTGGGTTGAGGACATCCAGATGGCCATTGACCTGGCGGAGAGAGCAGCC 642
      |||
      |||
      |||
2639 CGGCGCTGAGTTCTGGCCAGCAGCGCCCTGACAAAGTCCCTGATGAAGCCAGC 2698
      |||
      |||
      |||
641 CGGCGCTGAGTTCTGGCCAGCAGCGCCCTGACAAAGTCCCTGATGAAGCCAGC 582
      |||
      |||
      |||
2699 CGGCTGACCAAGAGTGAAGATGACCTGAGCGCTCGGCGGACATGGCTGAGCGCAG 2758
      |||
      |||
      |||
581 CGGCTGACCAAGAGTGAAGATGACCTGAGCGCTCGGCGGACATGGCTGAGCGCAG 522
      |||
      |||
      |||
2759 CCGCGCAGCGGCGGCAACAATGATGATGATGATGATGATGATGATGATGATGATGAT 2818
      |||
      |||
      |||
521 CCGCGCAGCGGCGGCAACAATGATGATGATGATGATGATGATGATGATGATGATGAT 462
      |||
      |||
      |||
2819 TGGTGGATTGAGCATCGAGTGGAGAAATCAAGTTGTTCTGAAAACCTGCTGAGGAAATTTCA 2878
      |||
      |||
      |||
461 TGGTGGATTGAGCATCGAGTGGAGAAATCAAGTTGTTCTGAAAACCTGCTGAGGAAATTTCA 402
      |||
      |||
      |||
2879 AAAACAGCAACGGGTGGCAAGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2938
      |||
      |||
      |||
401 AAAACAGCAACGGGTGGCAAGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 342
      |||
      |||
      |||
2939 ACAAAATCAGCAGCAGCATCATCCCTTGGCAGCTGCTGCTCGGCTACTGCTCA 2998
      |||
      |||
      |||
341 ACAAAATCAGCAGCAGCATCATCCCTTGGCAGCTGCTGCTCGGCTACTGCTCA 282
      |||
      |||
      |||
2999 CCATCCCTCTGAGTCCAGAAACATCCAGAAAGCTAGCTGTTCAAGTCTCACTTCAAGT 3058
      |||
      |||
      |||
```

```
Db 281 CCATCCCTCTGAGTCCAGGACATCCAGAAAGCTAGCTGTTCAAGTCTCAAGT 222
Qy 3059 CCCAGCTCTACTACTTTCAGGGCGGAAAGCAGTACAGTTTGAAGTGAAGTGA 3118
      |||
      |||
      |||
Db 221 CCAGCTCTACTACTTTCAGGGCGGAAAGCAGTACAGTTTGAAGTGAAGTGA 162
      |||
      |||
      |||
Qy 3119 TCCGAGTCCACACAGCTCTGCTGCGAGCCACAGTGTGAGCCACAAAGTCTCTTG 3178
      |||
      |||
      |||
Db 161 TCCGAGTCCACACAGCTCTGCTGCGAGCCACAGTGTGAGTCACAAAGTCTCTTG 102
      |||
      |||
      |||
Qy 3179 TGTATTGATGCGCGGACACACTCGTTTCCGAGTGGCTGTTCTTCTGAAAGCTTTCT 3238
      |||
      |||
      |||
Db 101 TGTATTGATGCGCGGACACACTCGTTTCCGAGTGGCTGTTCTTCTGAAAGCTTTCT 42
      |||
      |||
      |||
Qy 3239 TTCTTCTGTTTAAATGA-AGCTGTGTAATTAACACCTGT 3278
      |||
      |||
      |||
Db 41 TTCTTCTGTTTAAATGAATGTAATTAACACCTGT 1
      |||
      |||
      |||
```

RESULT 10

BX364605/c

LOCUS

BX364605 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens

DEFINITION

CDNA clone CS0DC027YA10 3-PRIME, mRNA sequence.

ACCESSION

BX364605

VERSION

BX364605.2 GI:46287877

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

REFERENCE

Li W.B., Gruber, C., Jessee, J. and Poyates, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 5, 2003 this sequence version replaced gi:30368816.

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

7229.f

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?c=CS1AC0072E06NP1&c=7229.f>.

Location/Qualifiers

1..951

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DC027YA10"

/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

```
Query Match 26.2%; Score 900.4; DB 4; Length 951;
Best Local Similarity 99.1%; Pred. No. 7.5e-230;
Matches 945; Conservative 2; Mismatches 3; Indels 4; Gaps 4;
```

```
2323 GAGTTCATCCGCTCGGGGAGGCTCAGCAAGCTCTCGGGGAAGGGCTCCAGCAGCGCATG 2382
      |||
      |||
      |||
951 GAGTTCATCCGCTCGGGGAGGCTCAGCAAGCTCTCGGGGAAGGGCTCCAGCAGCGCATG 893
      |||
      |||
      |||
2383 TTCTTCTGTTCAACGAGCTCTGCTATACAGAGCGGGGCTGACGAGCTTCAATGATG 2442
      |||
      |||
      |||
```

Db	892	TTCTTTCCTGTTTCAACGAGCTGCTGTGTATTAACAAGAGCCGGGGGCTGACGGCCCTCAATCAAG	833
Qy	2443	TTTAAAGTCCACGGGCAAGCTCCCGCTCTATGTGCATGACGATTTAGAGAGCGAAGACGAG	2502
Db	832	TTTAAAGTCCACGGGCA - STCCCGCTTATGGCAAGATTTGAGGAGCGAAGACGAG	774
Qy	2503	TGGGGGGGGGCCCACTGCTGACCCCTCGGGGGCCAGGGGCAAGTCCATCATGTGTGGCCGC	2562
Db	773	T-TGGGGGGGGCCCACTGCTGACCCCTCGGGGGCCAGGGGCAAGTCCATCATGTGTGGCCGC	715
Qy	2563	AGTTCTCGGTCGAGATGAGAGAAGTGGTTGAGGACATCCAGATGGCAATTGACCTGTGCG	2622
Db	714	AGTTCTCGGTCGAGATGAGAGAAGTGGTTGAGGACATCCAGATGGCAATTGACCTGTGCG	655
Qy	2623	GAGAAAGACAGACAGCCCGCCCTCTGATGTTCTTGAGCAGACGCCCTTGACAAACAATGCC	2682
Db	654	GAGAAAGACAGACAGCCCGCCCTCTGATGTTCTTGAGCAGACGCCCTCTGACAAACAATGCC	595
Qy	2683	CCCTATTAAGCCACCGCGGCTGACCAAGAGTCAAGAGATTGACTTGAGGGCTTCGGGCA	2742
Db	594	CCCTATTAAGCCACCGCGGCTGACCAAGAGTCAAGAGATTGACTTGAGGGCTTCGGGCA	535
Qy	2743	TGCGTGGAGCCGACAGGCCCGGCAACCGCGGCAACAATATGATGACAGTGTGTCGACCGC	2802
Db	534	TGCGTGGAGCCGACAGGCCCGGCAACCGCGGCAACAATATGATGACAGTGTGTCGACCGC	475
Qy	2803	AACACCAAGCGTCTCCATGTGTGAATTCCAGATCCGATGTGAAGATCAATTGCTTGGAAAC	2862
Db	474	AACACCAAGCGTCTCCATGTGTGAATTCCAGATCCGATGTGAAGATCAATTGCTTGGAAAC	415
Qy	2863	CTGCTGAGGAAATTTCAAAAACAGCAAGGGGTGGCAAGGCTGTGGGTGTGTTCACAAAAC	2922
Db	414	CTGCTGAGGAAATTTCAAAAACAGCAAGGGGTGGCAAGGCTGTGGGTGTGTTCACAAAAC	355
Qy	2923	TTCTGCGCTGTTCTTCTTAACAATCAACAAGGACAATCACTCCCTTGACAGGCTGCTCTG	2982
Db	354	TTCTGCGCTGTTCTTCTTAACAATCAACAAGGACAATCACTCCCTTGACAGGCTGCTCTG	295
Qy	2983	CTCGGCTACTTCGCTACACCATCCCTCTGTAGTCCGAGAACATCCAGAAAGACTACGTGTTC	3042
Db	294	CTCGGCTACTTCGCTACACCATCCCTCTGTAGTCCGAGAACATCCAGAAAGACTACGTGTTC	235
Qy	3043	AAGCTGCACTTCAAGTCCCAAGCTCTACTTACAGGCGGAAAGCGAGTACACGTTGCA	3102
Db	234	AAGCTGCACTTCAAGTCCCAAGCTCTACTTACAGGCGGAAAGCGAGTACACGTTGCA	175
Qy	3103	AGGTGAGATGGAAGATGCCGAGTGCACCAAGCTCTGCTGGCAACCCCAAGTGTAGC	3162
Db	174	AGGTGAGATGGAAGATGCCGAGTGCACCAAGCTCTGCTGGCAACCCCAAGTGTAGT	115
Qy	3163	CACAAAGAGTCTCTGTGTATTTGATATGGCCGAGACACTCGTTTT--CCGAGATGGGCTGCTTT	3221
Db	114	CACAAAGAGTCTCTGTGTATTTGATATGGCCGAGACACTCGTTTTCCGAGATGGGCTGCTTT	55
Qy	3222	CTTGGAAGAGCTTTCTTTCTTCTGTATTAAGAACTGTGTAATAATTACACC	3275
Db	54	CTTGGAAGAGAGTTCCTTTCTTCTGTATTAAGAACTGTGTAATAATTACACC	1
RESULT 11			
BQ881649			
LOCUS			
DEFINITION			
BO881649			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
1 (bases 1 to 912)			
BO881649			
AGNCOCURT 8728249 NIH_MGC_47			
5', mRNA Bequence.			
BO881649			
BO881649.1 GI:22273657			
EST.			
Homo sapiens (human)			
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Homnidae; Homo.			
1 (bases 1 to 912)			

FEATURES	source
<p>AUTHORS NIH-MGC http://mgc.nci.nih.gov/. TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished (1999) COMMENT Contact: Robert Strausberg, Ph.D. Email: cgsbbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LNCM253 row: n column: 12 High quality sequence stop: 684. Location/Qualifiers</p>	<p>1. .912 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6339275" /tissue_type="neuroblastoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 47" /note="Organ: brain; Vector: pOT8; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-AT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."</p>
<p>ORIGIN</p>	
<p>Query Match 25.7%; Score 883.4; DB 3; Length 912; Best Local Similarity 98.7%; Pred. No. 2.7e-225; Matches 901; Conservative 0; Mismatches 11; Indels 1; Gaps 1</p>	
<p>592 AAAAAATAATCATACCTCAGCAGAGCGCACTAGAGGACAAATCGTGAATTTCCACCAT 651 1 AAAAAATAATCATACCTCAGCAGAGCGCACTAGAGGACAAATCGTGAATTTCCACCAT 60</p>	
<p>652 AACCAATTGGACAAACACGACGAGATCAGATTTCCAGCTCTAGAGATTGCCCGTCGG 711 61 AACCAATTGGACAAACACGACGAGATCAGATTTCCAGCTCTAGAGATTGCCCGTCGG 120</p>	
<p>712 CTAGAGATGTATGGAATCCGGTTCCACCGGCCAAGACAGGAAAGCAGAAATCAAT 771 121 CTAGAGATGTATGGAATCCGGTTCCACCGGCCAAGACAGGAAAGCAGAAATCAAT 180</p>	
<p>772 CTGGCCGTTGGCAACAGGGAATTCATAGTTTCAGGGTTTCACCTAAGATCAATGCCCTTC 831 181 CTGGCCGTTGGCAACAGGGAATTCATAGTTTCAGGGTTTCACCTAAGATCAATGCCCTTC 240</p>	
<p>832 AACTGGGCCAAGGTGCGGAAGCTCAGCTTCAAGAGAAAGCGCTTTTCATCAAGCTCCGG 891 241 AACTGGGCCAAGGTGCGGAAGCTCAGCTTCAAGAGAAAGCGCTTTTCATCAAGCTCCGG 300</p>	
<p>892 CCAATATGCCAATATGCGGTACACAGATATCCTTGGAAATTCCTGAATGCGCAGTCGGGATTC 951 301 CCAATATGCCAATATGCGGTACACAGATATCCTTGGAAATTCCTGAATGCGCAGTCGGGATTC 360</p>	
<p>952 TGCAAGTCCTTCTGGAAATATGTGTGTAACATCATGCTCTTTAGACTTTTGAAGAG 1011 361 TGCAAGTCCTTCTGGAAATATGTGTGTAACATCATGCTCTTTAGACTTTTGAAGAG 420</p>	
<p>1012 CCCAACCAGAACCCCAAGCCCGTCTCTTTAGCCGCGGGGTATCATTTCCGTTCAATGCT 1071 421 CCCAACCAGAACCCCAAGCCCGTCTCTTTAGCCGCGGGGTATCATTTCCGTTCAATGCT 480</p>	
<p>1072 CGGACTAGAAAGCAGGTCTTCGATATGTTTAAAGAGAGACATTAAGAGTGCATGTT 1131 481 CGGACTAGAAAGCAGGTCTTCGATATGTTTAAAGAGAGACATTAAGAGTGCATGTT 540</p>	

Qy	1132	GAAGAAGACACACCAAGATTCATTCTATCCGGAGCCTTGCTGTACACGCTTACAGAACTG	1131
Db	541	GAAGAAGAACACACCAAGATTCATTCTATCCGGAGCCTTGCTGTACACGCTTACAGAACTG	600
Qy	1192	AATTCGAGAGTCGTGAGCAGTCTCAGCAGACACACAGCTTATCATTTGTGAGAAAGTGC	1251
Db	601	AATTCGAGAGTCGTGAGCAGTCTCAGCAGACACACAGCTTATCATTTGTGAGAAAGTGC	660
Qy	1252	GAATCTCCAGGGGGCCAGAGCTGCGCGGCGAGAAAGGAACCGAAGTTTCCGCCGGGGAG	1311
Db	661	GAATCTCCAGGGGGCCAGAGCTGCGCGGCGAGAAAGGAACCGAAGTTTCCGCCGGGGAG	720
Qy	1312	CCGGGGTCGCACCCGAGGCTTGCCCTCGAGAGAAAGCCCCGGGGTAAACAACAGGCGAC	1371
Db	721	CCGGGGTCGCACCCGAGGCTTGCCCTCGAGAGAAAGCCCCGGGGTAAACAACAGGCGAC	780
Qy	1372	GGAGCCGCGCTCGGCGGCCACGAGAGAAAGAGAGGTCTTTAAGATAGAGACCGACAG	1431
Db	781	GGAGCCGCGCTTGCGGCGGCCACGAGAGAAAGAGAGGTCTTTAAGATAGAGACCGACAG	840
Qy	1432	AGTAAACCTCAGCCCCCGCAGCCAAAGACAGAGTTCCTGACTGGCAGTCTTACCTTTCC	1491
Db	841	AGTAAACCTCAGCCCCCGCAGCCAAAGACAGAGTTCCTGACTGGCAGT-CTACCTTTTC	899
Qy	1492	GAGCTGTCTGTGA	1504
Db	900	GAGCTGTCTGTGA	912

RESULT 12			
BU931960			
LOCUS	904 bp	mRNA	linear
DEFINITION	AGGCCCTT 10485586 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:6672254 5', mRNA sequence.		EST 18-OCT-2002

ACCESSION	BU931960
VERSION	BU931960.1
KEYWORDS	GI:24120779
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Homo sapiens

REFERENCE	1 (bases 1 to 904)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

```

Tissue: Agouti-2002-2
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LRCM2952 row: h column: 14
High quality sequence stop: 721.
Location/Qualifiers
1..904
FEATURES
source
```

```

source
1..904
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6672254"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_idb="NIH_MGC_109"
/note="Organ: ovary; Vector: pOTB1; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGCGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald W. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

```

ORIGIN			
Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."			
Query Match	25.4%	Score 873.2;	DB 3; Length 904;
Best Local Similarity	98.6%	Prod. No. 1.5e-222;	
Matches 891; Conservative	0;	Mismatches 10;	Indels 3; Gaps 1;

QY	1015	AAACCAAAGCCAAAGCCCGCTCTCTTTAGACGGGGGCAATCATTTCCGTTCAAGTGCCG	1074
Db	1	AAACCAAAGCCAAAGCCCGCTCTCTTTAGACGGGGGCAATCATTTCCGTTCAAGTGCCG	60
QY	1075	ACTCAGAAAGAGTTCGACTATGTTTAAAGAGAGACATAGAAGTGCGAGTTTGAA	1134
Db	61	ACTCAGAAAGAGTTCGACTATGTTTAAAGAGAGAGCATAGAAGTGCGAGTTTGAA	120
QY	1135	AGGAAGACAGCAAGATTCTATTCCTAACGGCTTCTTCAAGCCTTACAGAACTGAA	1194
Db	121	AGGAAGACAGCAAGATTCTATTCCTAACGGCTTCTTCAAGCCTTACAGAACTGAA	180
QY	1195	TCCGAAGTGTGGAGCAGTCTCAGCAGAGACCAAGCTTACATTTTGGAGAAAGTCCGAA	1254
Db	181	TCCGAAGTGTGGAGCAGTCTCAGCAGAGACCAAGCTTACATTTTGGAGAAAGTCCGAA	240
QY	1255	TCTCCAGGGGCCAGAGCTGCGCGCAGAGAAAGAAACCGAAGTTTCCGCGGGAGCCG	1314
Db	241	TCTCCAGGGGCCAGAGCTGCGCGCAGAGAAAGAAACCGAAGTTTCCGCGGGAGCCG	300
QY	1315	GGGTCCGACCCGAGCCTCTGCGCCGAGAGAAAGCCCGCGGGTTAACAGCAGGCGGACGGA	1374
Db	301	GGGTCCGACCCGAGCCTCTGCGCGCAGAGAAAGCCCGCGGGTTAACAGCAGGCGGACGGA	360
QY	1375	GCGGCTCCGGCGCCACGAGAGAAAGAGAGAGTGCCTTAAGAGTATGAGACCCAGCAGAT	1434
Db	361	GCGGCTCCGGCGCCACGAGAGAAAGAGAGAGTGCCTTAAGAGTATGAGACCCAGCAGAT	420
QY	1435	AAACCTCAGCCCCCGCAGCCAGCAGACAGAGCTCCCTGACCTGGCAGTCTCACTTTCCGAG	1494
Db	421	AAACCTCAGCCCCCGCAGCCAGCAGACAGAGCTCCCTGACCTGGCAGTCTCACTTTCCGAG	480
QY	1495	CTGTCTGTGAATTCGCAAGGGGGAGTGGCCCCCTGCAACGTGAACCTTGTCTCCCAACTG	1554
Db	481	CTGTCTGTGAATTCGCAAGGGGGAGTGGCCCCCTGCAACGTGAACCTTGTCTCCCAACTG	540
QY	1555	AGCCCCGAGCAACAAAGCAGGCTCTCCCTTGAATCAGCCCGGTGATGAATGACAGGCTGC	1614
Db	541	AGCCCCGAGCAACAAAGCAGGCTCTCCCTTGAATGATGACCCGCTGTAATGACAGGCTGC	600
QY	1615	CCCCGAGCAGACGATGAGATGAGGGCCGAGAGAAAGATTCCCACTGATPAAAGCGTAC	1674
Db	601	CCCCGAGCAGACGATGAGATGAGGGCCGAGAGAAAGATTCCCACTGATPAAAGCGTAC	660
QY	1675	TTCTATAGCTAAGAAAGTGTCTACACCGAGCGAAATATTGAAAGATCTTCGAAAGTTATC	1734
Db	661	TTCTATAGCTAAGAAAGTGTCTACACCGAGCGAAATATTGAAAGATCTTCGAAAGTTATC	720
QY	1735	ACTTCGGGTTTGAAGCAAGTGAAGCAAGAAAGAGAGCCATGCGGAGACCTGAAGAAAGT	1794
Db	721	ACTTCGGGTTTGAAGCAAGTGAAGCAAGAAAGAGAGCCATGCGGAGACCTGAAGAAAGT	780
QY	1795	CTCATATTTCCGAATTTTGAACCTTTGCACAAATTTTATCTAATTTTCTCAAGGAATTT	1854
Db	781	CTCATATTTCCGAATTTTGAACCTTTGCACAAATTTTATCTAATTTTCTCAAGGAATTT	840
QY	1855	GAGCAACGACTTGCCCTCTGTTGGAGAGGCGGCTCAAAATGCCAAATCAGAGATTACAA	1911
Db	841	GAGCAACGACTTGCCCTCTGTTGGAGAGGCGGCTCAAAATGCCAAATCAGAGATTACAA	900
QY	1912	AGAA 1915	
Db	901	NGAA 904	


```

RESULT 13
BM551766 1026 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6575294 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5479350
DEFINITION 5', mRNA sequence.
ACCESSION BM551766
VERSION BM551766.1 GI:18789071
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 1026)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.jnl.gov
Plate: LHCN2000 row: h column: 07
High quality sequence stop: 675.
Location/Qualifiers
1. 1026
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5479350"
/issue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 98"
/note="Organ: brain; Vector: pORF7, Site 1, XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 25.1%; Score 863.8; DB 2; Length 1026;
Best Local Similarity 96.7%; Pred. No. 5,2e-220;
Matches 916; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

QY 2045 GAGACTTTGAGCTGACAGAGGTGTGTACCTACCGCTCAACACCTTCTCTCGGCGCAC 2104
DB 1 GAGACTTTGAGCTGACAGAGGTGTGTACCTACCGCTCAACACCTTCTCTCGGCGCAC 60
QY 2105 TGCACCGGCTCATGCACTACAGACAGCTCTGAGAGGGCTGTGTGCAAAACACACCCGCCCA 2164
DB 61 TGCACCGGCTCATGCACTACAGACAGCTCTGAGAGGGCTGTGTGCAAAACACACCCGCCCA 120
QY 2165 GCCACGCCGACTTCAGGAGCTGCCGAGCCGCTTTGGCAGAGATCAGAGATGTTGGAC 2224
DB 121 GCCACGCCGACTTCAGGAGCTGCCGAGCCGCTTTGGCAGAGATCAGAGATGTTGGAC 180
QY 2225 AGCTCCAGGTACGATGATCAAGATGAGAGATTTCAGAAAGTGCAGAACTCAAGAAAG 2284
DB 181 AGCTCCAGGTACGATGATCAAGATGAGAGATTTCAGAAAGTGCAGAACTCAAGAAAG 240
QY 2285 ATTGATTTGGCATTTGCAATCTTGTTCCGGGAGAGGAGTCAATCCGCTGGGAGCC 2344
DB 241 ATTGATTTGGCATTTGCAATCTTGTTCCGGGAGAGGAGTCAATCCGCTGGGAGCC 300
QY 2345 TCAGCAAGCTCTCGGGGAGAGGGGCTCAGACAGCGCATGTTCTTCTTCAAGAGTCC 2404

```

```

DB 301 TCAGCAAGCTCTCGGGGAGAGGGGCTCAGACAGCGCATGTTCTTCTTCAAGAGTCC 360
QY 2405 TGTATATACAGAGCCGGGGCTGACCGCTTCCAAATAGTTTAAAGTCCAGGGCAGCTCC 2464
DB 361 TGTATATACAGAGCCGGGGCTGACCGCTTCCAAATAGTTTAAAGTCCAGGGCAGCTCC 420
QY 2465 CGCTCTATGAGCATGAGATTTAGAGAGAGGAGAGACAGAGTGGGGGTGCCCACTGCTGA 2524
DB 421 CGCTCTATGAGCATGAGATTTAGAGAGAGGAGAGACAGAGTGGGGGTGCCCACTGCTGA 480
QY 2525 CCCTCCGGGGCCAGCGGCGAGTCCATCATCTGTGGCCGCCAGTTCTCGTCCGAGATGAGA 2584
DB 481 CCCTCCGGGGCCAGCGGCGAGTCCATCATCTGTGGCCGCCAGTTCTCGTCCGAGATGAGA 540
QY 2585 AGTGGTTTGAAGACATCCAGATGAGCCATGACTGGCCGAGAGAGACAGACAGCCGCCCC 2644
DB 541 AGTGGTTTGAAGACATCCAGATGAGCCATGACTGGCCGAGAGAGACAGACAGCCGCCCC 600
QY 2645 CTGAGTTCTTGCCGACAGACAGCCCTTGACACAAAGTCCCTGTATGAAAGCCAGCCGCTG 2704
DB 601 CTGAGTTCTTGCCGACAGACAGCCCTTGACACAAAGTCCCTGTATGAAAGCCAGCCGCTG 660
QY 2705 ACCAGAGTCAAGAGATGACCTGAGCGCTCGGACATGCTGAGCGCCAGCGCCCGC 2764
DB 661 ACCAGAGTCAAGAGATGACCTGAGCGCTCGGACATGCTGAGCGCCAGCGCCCGC 720
QY 2765 ACCGCGGCAACACATGATGTGACGTGTGTCGACCGCAACACAGCTCTCCATGATG 2824
DB 721 ACCGCGGCAACACATGATGTGACGTGTGTCGACCGCAACACAGCTCTCCATGATG 780
QY 2825 ACTTCAGATGCGAGTGAAGATCACTGTCTGAAACCTGCTGAGAAATTCAAAAACA 2884
DB 781 ACTTCAGATGCGAGTGAAGATCACTGTCTGAAACCTGCTGAGAAATTCAAAAACA 840
QY 2885 GCAAC-GGGTGGCAGAGCTGTGGGTGTGTTCACAACTTCTGCTG-----TTCCTT 2938
DB 841 GCAACGGGGTGGCAGAGCTGTGGGTGTGTTCACAACTTCTGCTGCTGTTCTTCT 900
QY 2939 ACAAATCACACCAAGACAT---CATCCCTTGGCCAGCTGCTGCTG 2982
DB 901 ACAAATCACACCAAGACAT---CATCCCTTGGCCAGCTGCTGCTG 947

RESULT 14
AL566821/c 930 bp mRNA linear EST 05-APR-2004
LOCUS AL566821 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF025YJ01.3-PRIME, mRNA sequence.
ACCESSION AL566821
VERSION AL566821.1 GI:12919564
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 930)
Li, W.-B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segre@genoscope.cns.fr; Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 7229.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna78=CS0DF025CE01NP1c=7229.f.
Location/Qualifiers

```

```

source
1. .930
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF025J01"
/issue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_1b="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (d)T primer. Five prime end
enriched, double-strand cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match      25.0%; Score 862; DB 1; Length 930;
Best Local Similarity 98.2%; Pred. No. 1.5e-219;
Matches 915; Conservative 0; Mismatches 10; Indels 7; Gaps 4;

OY 2366 GGCTCCAGCAGCGCATGTTCTTCTGTTCAACGACGTCCTGCTATACAGACCGGGGGC 2425
DB 930 GGCTCCAGCAGCGCATGTTCTTCTGTTCAACGACGTCCTGCTATACAGACCGGGGGC 871
OY 2426 TGAACGCGCTCCATCATGTTTAAAGTCACGGGCACTCCGCTCTATGCGATGACGATTG 2485
DB 870 TGAACGCGCTCCATCATGTTTAAAGTCACGGGCACTCCGCTCTATGCGATGACGATTG 811
OY 2486 AGGAGGCGGAACGAGCGTGGGGTCCCACTGCTGACCTCCGCGGGCCAGCGGCACT 2545
DB 810 AGGAGGCGGAACGAGCGTGGGGTCCCACTGCTGACCTCCGCGGGCCAGCGGCACT 752
OY 2546 CCATCATCGTGGCGCGCATGTTCTCGTCCGAGTGAAGTGGGTTGAGCATCCAGA 2605
DB 751 CCATCATCGTGGCGCGCATGTTCTCGTCCGAGTGAAGTGGGTTGAGCATCCAGA 692
OY 2606 TGGCCATTGACCTGCGGAGAGAGAGACAGACGCCGCCCTGATGTTCTGCGGACAGCC 2665
DB 691 TGGCCATTGACCTGCGGAGAGAGAGACAGACGCCGCCCTGATGTTCTGCGGACAGCC 632
OY 2666 CCCCTGACAAAGTCCCTGATGAAGCCACCGCGGCTACACAGAGTGAAGAGATGACC 2725
DB 631 CCCCTGACAAAGTCCCTGATGAAGCCACCGCGGCTACACAGAGTGAAGAGATGACC 572
OY 2726 TGAGCGCCTCGGCAATCGCTGAGAGCGCCGCCGACCGCGCAACACATGATGTC 2785
DB 571 TGAGCGCCTCGGCAATCGCTGAGAGCGCCGACG-CCGCAACCGCGCAACATGATGTC 513
OY 2786 ACGTGTGTCGACCGCAACACCGGCTCTCCATGATGATGATCA--GCATGCACTGG 2842
DB 512 ACGTGTGTCGACCGCAACACCGGCTCTCCATGATGATGATCAACGCMCCGCGCACTGG 453
OY 2843 AGAATAGTGTCTGGAACCTGCTGAGAAATTCAAAACAGCAAGGAGGCGAGAAC 2902
DB 452 AGAATAGTGTCTGGAACCTGCTGAGAAATTCAAAACAGCAAGGAGGCGAGAAC 393
OY 2903 TGTG--GATGATGTTCAAACTTCTGCTGATCTTCTCAATCAACACAGGACATCA 2960
DB 392 TGTGAGGAGGATGTTCAAACTTCTGCTGATCTTCTCAATCAACACAGGACATCA 333
OY 2961 TCCCTTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3020
DB 332 TCCCTTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 273
OY 3021 CATTCAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 3080
DB 272 CATTCAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 213
OY 3081 GGAAGAGCGATCACTTCGAAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAAT 3140
DB 212 GGAAGAGCGATCACTTCGAAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAAT 153
OY 3141 CTGCGAAGCCCACTGTTGAGCAAAAGATCTTTGTGTATGATGCGCGGACCACT 3200
DB 3141 CTGCGAAGCCCACTGTTGAGCAAAAGATCTTTGTGTATGATGCGCGGACCACT 3200

```

```

DB 152 CTGCGAAGCCCACTGTTGAGTACAAAGATCTCTGTGTATGATGAGCCGAGACACT 93
OY 3201 CGTTCCGAGAGTGGCTGCTTCTGGAAGACGTTCTTCTGTTATTAAGAGCT 3260
DB 92 CGTTCCGAGAGTGGCTGCTTCTGGAAGACGTTCTTCTGTTATTAAGAGCT 33
OY 3261 GGTAAATTAACACCTGTCTGAAATCAAAA 3292
DB 32 GGTAAATTAACACCTGTCTGAAATCAAAA 1

RESULT 15
AY415515
LOCUS
DEFINITION
AY415515 Homo sapiens FARP2 gene, VIRUAL TRANSCRIPT, partial sequence,
AY415515 genomic survey sequence.
ACCESSION
AY415515 GI:39771474
VERSION
AY415515.1 GI:39771474
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 3165)
REFERENCE
Clark,A.G., Gianoweki,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanendbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Snihsy,J.J.,
Adams,M.D. and Cargill,M.
Interfering nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
AUTHORS
Clark,A.G., Gianoweki,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanendbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Snihsy,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
TITLE
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
location/Qualifiers
source
1..3165
/mol_type="Homo sapiens"
/db_xref="taxon:9606"
<1..>3165
/gene="FARP2"
/locus_tag="HCM5583"

ORIGIN
Query Match      24.5%; Score 842.4; DB 14; Length 3165;
Best Local Similarity 51.9%; Pred. No. 4.1e-214;
Matches 1606; Conservative 0; Mismatches 144; Indels 45; Gaps 7;

OY 69 GCCGACCCAGGATGACGATGCGGGGCCCGGAAATTCGGGATGACGATCTGGAACG 128
DB 33 GCGAGCTGCGAGGATGCGGCTTGGGTGCCAGACCCCTGTGGAGTTAGCACTTGAGCC 92
OY 129 TGAAGAGAGCCGCCCAACACCTTCAGAAAATCTGTGTCATCAAAATCAGATGCT 188
DB 93 TGGGAGAGCTGCTTGGCCAGAAATGCAAGAGACCTGCACTCAGATGAAGTGTCT 152
OY 189 GATAGACCCAGAGAGGATTTGAAGTTCCACAAAGAGCTCTGGAGAGTGTCTGTGA 248
DB 153 GGAACAACCACTAGAAATTTGACATTGAGCTTAATGCGATGCGCAGATTTACTGAC 212
OY 249 TGCAGTTGCAACCACTCACTGCTGGAAGAGTGAATTTGGCTGAGATTTCTCTGA 308
DB 213 ACAAGTGGAGAGCGTTTAACTGTATGATGTACTTCTGGAGTGAAGTTTCAAAA 272

```

QY 309 TCACAAAGATGACGGTGTGCTGATCTCTAAACCCATTGTGAAAGATTAGAG 368
 Db 273 TACTCAGTCTACTGGATTGGCTTGAACTATGAACCATCTATTGGCAATACGAAG 332
 QY 369 GCCAAAGCAGCTTGTGTGTTAAGTTGTGTGAAATCTTTCCGCTGACACACAACT 428
 Db 333 GCCAAAGAAATGTGTGCTGCTGCTAGCTGTAATAATTTTTCACCTGATCTGGTACGT 392
 QY 429 CCAAGAAAGACTACAAAGGTACCTGTTCCGCTGAGTGAAGAGAACTTGGCTCAAG 488
 Db 393 ACAAGAAAGAAATATCAAGATATCTTGTTCCTTGCACTTAAGAGAACCTGCTGGAAGA 452
 QY 489 CAGGTGACGTGTAATGACACACAGCGAGCTCTTGATTTCAACATTTGCAATGTA 548
 Db 453 GCGTTTGACCTGTGCTGACACACAGGGGCCCTTCTCAAGTCCATCTCTGCAGTGGA 512
 QY 549 GATTGGGATTTTGTATGAGCCCTTGAACAGAGACCTTAGACAAAATAATACATACC 608
 Db 513 AATAGGATTTACGATGAAACGCTGACCGAGAGCACTCAAAAGTGAACGATATTGGCC 572
 QY 609 TCAGCAAGACGACTAGAGACAAATACTGTGAATTTCACTATTAACCACTTGGACAAAC 668
 Db 573 TGCGCAGACGACTGCTCTTGAAGAAATATCAATTCATCAGAAAGCAGTGGGCCAGAC 632
 QY 669 ACCAGAGAAATCAGATTTCCAGCTCTAGAGATTGCGGCTAGAGATGTATGGAAT 728
 Db 633 ACCTGTGAGTCGGAATTTCCAGGTGCTCGAAATTTGCTCGAAATTTGGAATGTACGGCAT 692
 QY 729 CCGGTTGCAACCGGCCAAGAGAGGACAGAAATCAATCTGGCCGTTGCCAACAC 788
 Db 693 CAGATTTTCATGCTTCTGACAGGGAGGAACCAAGATTCATGCGCAGATTTCCCAT 752
 QY 789 GGGAAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCTTTCACTGGGCCAAGTGGC 848
 Db 753 GGGTGTACTGTGTTCAGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 812
 QY 849 GAAGCTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGGCCAGATGCCAATAGTG 908
 Db 813 NNGAAC 872
 QY 909 GTAACAGATACCTTGAATTTCTGATGCGCACTCGGATTTTCTGCAAGTCTTTGGAA 968
 Db 873 TTACACGAGACACTTGAATTTTGTGGGTAGTAGAATGTAAAGAACTTTGGA 932
 QY 969 AATCTGTGTGAACATCAATGCTTTTGAATTTTGAAGAGCCCAACCAAGCCCA 1028
 Db 933 GATTGTGTGAGATATCAACTTTTGTAGACTTTTGAACCAACTTAAGCCAAACCA 992
 QY 1029 GCCCGTCTCTTAGCCGCGGGGTCACTTTCGTTCAAGTGTGCGACTCAGAAAGAGT 1088
 Db 993 AGCCGCTTCTTACGCCCGGGGCTCTCTTCAAGATACAGNNNNNNNNNNNNNNNNNN 1052
 QY 1089 TCTCGACTATGTTAAAGAGAGACATTAAGAGGTGCAAGTTTGAAGAGACAGCAA 1148
 Db 1053 NNAAGCAGCAA 1112
 QY 1149 GATTCAATTCTATCCGAGAGCTTGTCTCACAGCTTACAGAACTTAATTCGAAAGTGTGGA 1208
 Db 1113 GACCCACAGCTCC-----GTTGAGCTCTGACTGCAAGACCTTACCAAA 1154
 QY 1209 GAGGTCTCAGACAGACAGCAGCTTACATTTGAGAAAGTGCAGAAATTCACAGGGGCCA 1268
 Db 1155 ACAGAGACTCTATTTCCCGAAGGATTGAGAGACTCTGCTCTCCCATTTTCAGCGAATGC 1214
 QY 1269 GAGCTGCGCGAGAGAAAGAACCGAAGTTTCCGCCGGGAGCCGGGGTTCGACCCGAG 1328
 Db 1215 CTTTTCATCGCTCTCTCTCTCACTCTGCTCCCTCTGCTGCTCAGAGTTTAAAGACAG 1274
 QY 1329 CCGTGGCCGAGAGAGAGCCCGCGGGGTAAACAGCAGCGAGCGAGCCGCTCGCGGCC 1388
 Db 1275 CAGCACTCCCTTACAGATCCCAAGGTTTCTTACGTCAAGAGTCCAGCTGACAGAGGGG 1334
 QY 1389 CACGAGAGAGAGAGAGG---TCGTTAAGATTAGAACCCAGCAGAGTTAACTCTAGCC 1445

Db 1335 CAGTGGAGAGTGGCTGAGAGCCCCCGACACACCATGTGGCCCAACCCCTGGGGCCCCCGC 1394
 QY 1446 CCGCAGCCACAGACAGAGCTCCCTGACTGCGAGTCTTCACTTTCCGAGTGTGTGAA 1505
 Db 1395 ACTTCAGCCCTGTGTCAGAGCCTTTCACAGAAAGATCTCAGCCTTCTCCCTCAGCCGGAA 1454
 QY 1506 CTCGAGAGGGGGAGTGGCCCCCTGCAACGTGACTTGTCTCCCACTGAGCCCCCAGAC 1565
 Db 1455 GAGCCCCCTGAGTCTGAGCCCTGCAATTTCAAGTGCTT-----TGGGCCAGCTGA 1505
 QY 1566 CAAGAGGCTCTCTCCCTGATTCAGCCCGCTGTGAATGACCAAGGCTGCCCCCGAGCGGA 1625
 Db 1506 ACAGGGCTATATCCCACTCCTGAGCCCTGTGCTCAAGTATGTGGGGAGCCGGATGGA 1565
 QY 1626 CGATGAGATGAGGGCCGAGAGAAAGATTTCCACTGATTAAGCTTACTTCAATGCTAA 1685
 Db 1566 CTCGAG---GAGCCACAGACAAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1622
 QY 1686 GGAAGTGTACACCGAGCGAACAATATCTGAAGATCTCGAAGTTATCACTTCGTTG 1745
 Db 1623 NNTGTGTT 1682
 QY 1746 TCAGACACAGTGAAGCAAGAGAGCGCAATGCGGAAGCACTGAAAGTCTCATATTCC 1805
 Db 1683 CCGACCGCAGTGTGTAAGAGAGAGCCCATGCTGCGACTGTGATGACGCTGCTTTCTC 1742
 QY 1806 GAAATTTGAACCTTTGCACAAAATTTCACTAATTTTCTCAAGAAATTGAGCAAGACT 1865
 Db 1743 CAACATCGATCCCATCTATGATGATTCACAGAGGCTTCTGCGAGTGTGAGCAGAGCT 1802
 QY 1866 TGCCCTGTGGGAAGGCGCTCAATGCCCCAATCAAG---ATTACAAAGATTCGGCA 1922
 Db 1803 GGCATCTGGGAAGGGCCCTCCAAAGCCACAAAGGCAATCAAGATCGGGGA 1862
 QY 1923 TGTCTATGCTGAAGAACTTCAAGGCGATGAAGCACTGCGGCTCACCTGTGAAAGCAAG 1982
 Db 1863 CATCTGCTCAGGAACATGCGCAATTAAGAGTTTACAGCTATTTCCAAAGACATGA 1922
 QY 1983 CGAGCCTTGAAGAGCCCTGAGAGATGAATCAAGACTCCCGCGCTGAGAACTTCTG 2042
 Db 1923 CGAGGTCTTAACGAAGCTGAAAGGCTCAACAACTCTTAAGAGTTGAGGCACTGTA 1982
 QY 2043 CAGAGCTTGAAGTGAAGAGTGTGTTACCAACGCTCAACACTTCTCTGCGGCC 2102
 Db 1983 CAAAGTTTGAAGTGAAGAGTGTGCTTACTTCTCTCAACACTTCTCTGTTAGGCC 2042
 QY 2103 ACTGCAACCGGCTCATGACTACAGAGAGTCTGAGCGGCTGTGCAACACCAACCGCC 2162
 Db 2043 CATCAGCGGCTGTGCACTACCGCTGCTGTGCGGCCCTTATGCGGACATTTACAGCCC 2102
 QY 2163 GAGCAGCGCCGACTTCAAGGACTGCGGAGCCGCTTTGCAAGATCAAGAGATGTGTGC 2222
 Db 2103 CCGGACCATGACTAGCTGACTGACTGATGACCGCTGAAGCATCAAGAGGTGACCAAC 2162
 QY 2223 ACAGCTCCACGATGAGATGATCAAGATGAGAAATTTCCGAAGCTGCAAGACTCAAGAA 2282
 Db 2163 CACACTACAGCAACTTCTATCTGCTGAGAACTTCGCAAGACTTAAAGAGTGCAGGG 2222
 QY 2283 AGATTTGATTTGCAATGCAATTTGTGTTCCGAGAAAGGAGTTATCTGTTGCGAG 2342
 Db 2223 GGAACGTGTGGCATAGAAACCTCATTTCTCTGCAAGAGAGTTTATCTGTAGGGCTG 2282
 QY 2343 CCTCAGACAGTCTCGGGGAAGGGCTCAAGACGCGCATGTTCTTCTGTTCACAGCT 2402
 Db 2283 CTTTCAAGAGCTCACCAAGAAAGGCTCTGCAAGAGAGAGTGTTTTCTTCTCTCAGATAT 2342
 QY 2403 CCTGCTATACAGAGCCGGGGCTGACGCGCTTCCATTCAGTTTAAAGTCAAGGGAGCT 2462
 Db 2343 GTTGTGTACAAAGAAAGAGTTGCAAGAACCAAGCACTTCGGATTCGGGGCTTCT 2402
 QY 2463 CCGCTCTATGCAATGCAATTTGAGAGAGCGAAGCGAGTGGGGGTGCCCACTGCT 2522

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: May 22, 2006, 00:49:20 ; Search time 595.56 Seconds
(without alignments)
10813.957 Million cell updates/sec

Title: US-09-555-342B-1

Perfect score: 3442
Sequence: 1 CGCGCAGCGCGCGCGCGCTG.....AGTATCAACATGTCATC 3442

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

1: /EMC_Celerra_SIDS3/ptodata/2/1na/1_COMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/1na/5_COMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/1na/6_COMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/1na/7_COMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/1na/7_COMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/1na/7_COMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/1na/7_COMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/1na/7_COMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/1na/7_COMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/1na/7_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	238	6.9	310	3	US-09-621-976-8193 Sequence 8193, App
2	192.4	5.6	2853	3	US-09-949-016-1933 Sequence 1933, App
3	180.4	5.2	6263	3	US-09-664-958-5 Sequence 5, App1
4	180.4	5.2	6268	3	US-09-566-921-57 Sequence 57, App1
5	169.6	4.9	4292	3	US-09-949-016-1488 Sequence 1488, App
6	169.6	4.9	4336	3	US-09-949-016-208 Sequence 208, App
7	165.8	4.8	576	3	US-09-270-767-1995 Sequence 1995, App
8	165.8	4.8	576	3	US-09-270-767-1995 Sequence 1995, App
9	144	4.2	3984	3	US-09-848-294-1 Sequence 1, App1
10	143.2	4.2	2872	3	US-09-906-779-3 Sequence 3, App1
11	142.4	4.1	2156	3	US-09-949-016-3593 Sequence 3593, App
12	133.6	3.9	3398	3	US-09-799-451-201 Sequence 201, App
13	106.4	3.1	449	3	US-09-270-767-1542 Sequence 1542, App
14	106.4	3.1	449	3	US-09-270-767-1542 Sequence 1542, App
15	105.2	3.1	184	3	US-09-621-976-15728 Sequence 15728, App
16	97.8	2.8	17098	3	US-09-949-016-15864 Sequence 15864, App
17	95	2.8	2493	3	US-10-104-047-41 Sequence 41, App1
18	90.6	2.6	573	3	US-09-270-767-165 Sequence 165, App
19	90.6	2.6	573	3	US-09-270-767-165 Sequence 165, App
20	65.2	1.9	3438	3	US-10-164-595-29 Sequence 29, App1
21	63.8	1.9	4272	3	US-09-949-016-5747 Sequence 5747, App
22	63.8	1.9	1626	3	US-09-620-312D-1033 Sequence 1033, App
23	63.8	1.9	1699	5	US-09-913-171A-26 Sequence 26, App1

24	62.8	1.8	2730	3	US-09-774-528-175 Sequence 175, App
25	62.8	1.8	2730	3	US-10-120-988-175 Sequence 175, App
26	62.8	1.8	3503	4	US-10-094-749-560 Sequence 560, App
27	59.6	1.7	722	3	US-09-270-767-1861 Sequence 1861, App
28	59.6	1.7	722	3	US-09-270-767-1861 Sequence 1861, App
29	59.4	1.7	222	3	US-09-016-434-45 Sequence 45, App1
30	56.8	1.7	4081	2	US-08-446-345-35 Sequence 35, App1
31	56.8	1.7	4081	2	US-08-446-345-35 Sequence 35, App1
32	53.8	1.6	7218	2	US-08-232-463-14 Sequence 14, App1
33	50.2	1.5	3044	4	US-09-814-915A-84 Sequence 84, App1
34	50.2	1.5	3044	4	US-09-880-107-3718 Sequence 3718, App
35	48.8	1.4	1308	3	US-09-252-991A-11936 Sequence 11936, App
36	48.8	1.4	1335	3	US-09-252-991A-11936 Sequence 11936, App
37	48.2	1.4	1816	3	US-09-620-312D-439 Sequence 439, App
38	48.2	1.4	3415	3	US-09-902-540-5107 Sequence 5107, App
39	48.2	1.4	28058	3	US-09-902-540-5107 Sequence 5107, App
40	47.4	1.4	2400	3	US-08-930-001-1 Sequence 1, App1
41	47.4	1.4	2400	3	US-08-930-001-1 Sequence 1, App1
42	47	1.4	801	3	US-09-902-540-8879 Sequence 8879, App
43	47	1.4	12028	3	US-09-902-540-8879 Sequence 8879, App
44	47	1.4	54711	3	US-09-949-016-17489 Sequence 17489, App
45	46	1.3	235452	3	US-09-949-016-13675 Sequence 13675, App

ALIGNMENTS

RESULT 1
US-09-621-976-8193
Sequence 8193, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jober, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET 054PR2
CURRENT APPLICATION NUMBER: US/09/621, 976
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 8193
LENGTH: 310
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-8193

Query Match 6.9%; Score 238; DB 3; Length 310;
Best Local Similarity 95.3%; Pred. No. 1.6e-53;
Matches 244; Conservative 10; Mismatches 1; Indels 1; Gaps 1;

QY	1	CGCCGAGCGCGCGCGCT-GTGAAGATTTCTTAAGCGCTTTCATCATGGAGAAT	59
DB	55	CGCCGAGCGCGCGCGCTGTGAGAKTKCTTAAGCGCTTTCATCATGGAGAAT	114
QY	60	AAGAGAGAGCGAGCGCGCGCTGAGTACGAGTGGGGCGCGGAAATTCGGGATCAGTAC	119
DB	115	AAGAGAGAGCGAGCGCGCGCTGAGTACGAGTGGGGCGCGGAAATTCGGGATCAGTAC	174
QY	120	CTTGAAGCTGAGAGAGCGCGCGCTGAGTACGAGTGGGGCGCGGAAATTCGGGATCAGTAC	179
DB	175	CTTGAAGCTGAGAGAGCGCGCGCTGAGTACGAGTGGGGCGCGGAAATTCGGGATCAGTAC	234
QY	180	CGAGATGCTGATGACACCGAGAGGATTTAAGTTCAACAAGAGCTCTGGGAGGT	239
DB	235	CGATATGCTGATGACACCGAGAGGATTTAAGTTCAACAAGAGCTCTGGGAGGT	294
QY	240	GCTGCTGATGAGT 255	
DB	295	GCTGCTGATGAGT 310	

RESULT 2

US-09-949-016-1933
 ; Sequence 1933, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMOPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 1933
 ; LENGTH: 2853
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-1933

Query Match 5.6%; Score 192.4; DB 3; Length 2853;
 Best Local Similarity 52.8%; Pred. No. 7,8e-41;
 Matches 490; Conservative 0; Mismatches 426; Indels 12; Gaps 3;

165 CGTGTTCATCAAAATCCAGATGCTGGATACACCCAGGAGGCAATTTGAAGTCCCAAAAG 224
 784 CATGACATGCAAGGTTCTTTGTTGATGACACAGTTTATGATGTGTGAGAGAAACA 843
 225 AGCTCTGGGAGGAGTCTGCTGGATGCACTTTGGCAACCACTCAACTGTTGGAAGTGA 284
 844 TGCTAAGGAGCAAGATTTGCTTAAAGAGATATGAGCATCTCAATCTTTGGAAGAGA 903
 285 CTATTTTGGCTGAGTTTCTGATCACAAAAAGATCAGGTGTGGCTGGATCTCTTAA 344
 904 CTATTTTGGCTAGCATTTGGGATTAACCAACTCTTAAGACATGGCTGATTCGCCCA 963
 345 ACCCATTTGGAAACGATTAGAAAGGCAAGCAAGCTGTGTTTAAAGTTGTGTGTAAT 404
 964 AAAAAATAAAAAGAGGTTCGTGG---TGTCCTTGGAAATTTTACATTTAATGTAAAGTT 1020
 405 CTTTCGGCTGACCAACACATCCCAAGAAAGACTCAAGAGTACTGTTCGCGCTGCA 464
 1021 TTATCACTGACCAAGCAAGTAAACAGAAACATTAACATATTTATTTATGTCTTCA 1080
 465 GGTGAAGCAGACTTGGCTCAAGGCAAGTTGACGTGTAAATGACACCAAGCGACGCTCTT 524
 1081 GCTTCGGCAGGACATATGTGACAGACGTCTGCCCTGTTCCTTTGCAACCTTAGCATTAAT 1140
 525 GATTTCACACATTTGCAATCTGAGATTGGGGATTTTGAAGCTTGGACAGAGAGA 584
 1141 AGGTTCCTTACCAATCCAGTCTGAACTGGGAGACTACGACCAAGAACTCCATGGCGTGA 1200
 585 CTTAGCAAAAATTAATACATAC-----TCAGCAAGAGCAGCATAGAGCAAAATGCT 638
 1201 TTATGTAGTGAATTTTAACTGGCCCCGAATCAGACCAAGAACTTGAAGAGAGATCAT 1260
 639 GGAATTTCCATTAACACATTTGACAAACACACAGCAGATCAATTCAGCTCCTTGA 698
 1261 GGAATGCAATAGTACATAGCTCAGCTCAGCTCAGCTGACTGAGATTTTCTTGA 1320
 699 GATTCCCGCTGGCTAGAGATGTATGGAATCCGGTTGACCCGGGCAAGGAGAGGAGG 758
 1321 GAATGCCAAAAAGTGTCTATGTATGAGATGATCTTCAAAAGCAAGGACTTGAAGG 1380
 759 CACGAAGATCAATCTGCGCTTGGCAACACAGGAAATTCAGTGTTCAGGGTTTCACTA 818
 1381 AGTAATATATCAATCTGAGTGTCTGCTTAGTGCCCTTCGTGTTTCAAAAGATTAAGCTAG 1440
 819 GATCAATGCTTCAACTGGGCAAGGTGCGAAGAGCTGAGCTTCAAGAGAGAGCGCTTCT 878

1441 AATTAACCGCTTCCCTTGGCCCAAAAGTCTGAAGATTTCTTATAAAGTAGACTTTT 1500
 879 CATCAAGCTCGGCGCAAGATGCAATAGTCGTACAGAGATCTTGAAGTCTTATGAGC 938
 1501 CATCAAGATTCGGCTGGAGAGCAAGACAGATATGAAGTACATCGGATTTCAACTTCC 1560
 939 CAGTGGGATTTTCAGATCTCTTCTGGAATAATGTGTGAACATCATGCTCTCTTGA 998
 1561 CAGTTACGAGACAGTAAAGAAATTAATGAAAGTCTGTGTAGAACTACACAGTTTTCAG 1620
 999 ACTTTTGAAGAGCCCAACCAAGGCCAGCCGCTCTCTTTAGCCGGGGTCAATCATT 1058
 1621 A---TTGACATCTACAGACACCATTTCCCAAAAGCAAAATTTCTGGCTAGATCCAAAT 1677
 1059 TCGGTTGAGTGGTGGACTCAGAAAGCAG 1086
 1678 TCGATACAGTGGCGGACTCAAGCTCAG 1705

RESULT 3
 US-09-664-958-5

; Sequence 5, Application US/09664958
 ; Patent No. 6916912
 ; GENERAL INFORMATION:
 ; APPLICANT: Trakht, Ilya
 ; APPLICANT: Canfield, Robert
 ; APPLICANT: Kalantarov, Gary
 ; APPLICANT: Rudchenko, Sergei
 ; TITLE OF INVENTION: No. 6916912el Tumor-Associated Marker
 ; FILE REFERENCE: 0575/60240
 ; CURRENT APPLICATION NUMBER: US/09/664,958
 ; CURRENT FILING DATE: 2000-09-18
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 6263
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-664-958-5

Query Match 5.2%; Score 180.4; DB 3; Length 6263;
 Best Local Similarity 52.9%; Pred. No. 1.9e-37;
 Matches 485; Conservative 0; Mismatches 416; Indels 15; Gaps 4;

185 TGCTGATGACACCCAGGAGGCAATTTGAAGTTCCACAAAGAGCTCTGGAAGGTCTGC 244
 469 TGCTTGAATGCTCGAGATGATGATGTGAGTGGAGAAACATGCGCGGCAAGTGTCTGT 528
 245 TGGATGCAAGTTTGCAACCACTCAACTGTGGAAGGTGACTATTTTGGCTCGAGTTTC 304
 529 TTGACCTGCTGTGTAACCACTCAACCTCTAAGAAAGACTACTTGGCTGACCTTCT 588
 305 CTGATCACAAAAAGATCAAGGTGTGCTGATCTCTTAAACCCATTGTGAAACAGATTA 364
 589 GTGATGCTGACAGCCAGAAAGATGCGCTGAGACCCCTCCAAAGAGATCAAGAGATCC 648
 365 GAAGGCCAAAGCAAGTTGTGTTAAGTTTGGTGAATTTCTTTCGCGCTGACACACAGC 424
 649 GAG---TAGCCCTGGAATTTTGGCTTACAGTCAAGTTTCTACCGGCTATCTGTGCC 705
 425 AACTCCAAAGAACTCAAGAGTACCTGTTGCGGCTGACAGGTGAAGCAGGACTTGGCTC 484
 706 AGCTGACAGAAAGATCAACAGATACACTGTGCTGCTGACGCTGCGGAGACATCATCA 765
 485 AAGCAGGTTTACGTGTATATACACACAGCGAGCTCTTGTGATTTTCAACATTTGCAAT 544
 766 CGGCGCGGCTCAGATGCTCTTGTTCAGCATGCGCTTACTGAGCTCTTACCGTGTGAGG 825
 545 CTGAGATTTGGGATTTTGAAGCTTGAAGCTTGAACAGAGACCTTGAACAAAATAAATACA 604
 826 CTGAGCTGGGTGATATATATCTGAGAGCATGTGGGACATATATGTCAAGCAGACTCCGCT 885


```

RESULT 5
US-09-949-016-1488
; Sequence 1488, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```

/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 1488
/ LENGTH: 4292
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-1488

Query Match 4.9%; Score 169.6; DB 3; Length 4292;
Best Local Similarity 51.4%; Pred. No. 1.2e-34;
Matches 472; Conservative 0; Mismatches 434; Indels 12; Gaps 3;

```
175 AAAAAACAATGCTGATGACACCCGAGGAGCATTTGAAGTTCCAAAGAGCTCTGGG 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 675 AAGATGACCCCTCTTAAATGACACCGAATACAGCTGTGACCTGGAGAAACATGCCAAGGA 734

235 AAGTGTCTGTGATGACATTTGCAACCACTCAACCTGTGAGAGTACTATTTTGGC 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 735 CAAGTGTATTGACAAAGTGTGTGAACCTCAATCTTTGAGAAAGACTAATTGGA 794

295 CTCGAGTTTCTGTATCAAAAAAGATCAGGTGTGCTGATCTCTTAAACCATTTG 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 795 CTTTGTGTTTCAGGAAAGCCCTGACGAAAACTGTTAGATCTGCTAAAGAAATTAAG 854

355 AAACGATTAAGAAGCCAAAGACGTTGTGTTAAATTGTGTGAATCTTTCGGCT 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 855 AAGCAACTAGAAACC--TTCCATGCTATTCACTTTAAATGTGAAGTTTATCTCT 911

415 GACCAACAACAATCTCAAGAAAGACTCAAGATGACTTGTGCGCTGAGGTGAAGCAG 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 912 GATTCCTTCAATTGACTGAAGTATACCAAGATATCTTGTGCTTCAAGCTCGGAG 971

475 GACTTGTCTCAAGGAGGTTGACGTGTAAATGACACCAAGGACGCTCTTTGATTTCAAC 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 972 GACATGCTCTGCGCGCTGCGCTGCTGCTTTGTGATCATGCTCTCCGAGATCTTAC 1031

535 ATTGTGCATCTGAGATTGGGGATTT-----GATGAAGCCTTGAAGAGGACGACTTA 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1032 ACCCTGACAGGCTGAACTTGTGTGATGACCAAGAAACATGAGCAGCTCGACT 1091

589 GCAAAAAATTAATATACCTCAGCAAGACGCACTAGAGGACAAATCGTGAATTTTAC 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1092 GAATTCAGATTGGCCCTTACAGACTAAGAGCTGAAAGAGAGTGGCAGAGCTGAC 1151

649 CATTAACCAATTGACAAACACGACGAATCAGATTTCCAGCTCTTAAGATTGCCGT 708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1152 AAAACCCACAGGGGGCTTATCGCCAGCAGCAAGCTGATTCACAGTTCTTAAATAATGCAAG 1211

709 CGGCTAGAGATGATGGAATCGGGTTGCACCGGCCAAAGACGAGGAAGCAAGAAATC 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1212 AGGCTTTTCATGTATGTGTGATGACCTACATTCATGCAAGAGCTCAGAAAGTGTGACATC 1271

769 AATCTGCGCGTTGCCAACAAGGAATTTCTAATGTGTTTCAAGGGTTTCACTAAGATCAATGCC 828
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1272 AAGCTGGGGGTGTGTCTAATGGACTTCTCATTTTCAAAAGACGACTGGGAATCAATCGT 1331

829 TTCAACTGGGCCAAGGTGCGGAAGCTGACCTTCAAGAGAGAGCGCTTTCTCATCAAGCTC 888
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1332 TTTGCTTTGCGCAAAATCTTAAATAATTTCTTAAACGCAATCACTTCACTTAAAGTC 1391

889 CGGCCAGATGCAATAGTGCGTACAGAGTACCTTGGAATCTCTGATGGCCAGTGGCGAT 948
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1392 AAGCCGCAAGAGTGGAAACGTTTGAAGATCCATTGATTTCAAACTGCCAAACACCGG 1451

949 TTCTGCAAGTCTTCTGAAAAATCTGTGTGAAATCATATGCTCTTTTGAATTTTGA 1008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1452 GGAAGCAAAAGACTAAGAAAGTGTGTGAGCAATCATATCTTTCAAGGCTTGT--- 1508

1009 GAGCCCAAAACAAAGCCCAAGCCGCTCTTTAGCCGGGGGTCACTATTTGGTTCAAT 1068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 1509 TCTCCAGAGACGCCCAAAAAGCAAGTTCTGACCTTGGGGTCCAAATTTGCTAATG 1568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1069 GGTGGACTCAGAAAGCAG 1086
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1569 GGCCGACCCACAGACAG 1586
```

RESULT 6

US-09-949-016-208
Sequence 208, Application US/09949016

/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 208
/ LENGTH: 4336
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-208

Query Match 4.9%; Score 169.6; DB 3; Length 4336;
Best Local Similarity 51.4%; Pred. No. 1.2e-34;
Matches 472; Conservative 0; Mismatches 434; Indels 12; Gaps 3;

```
175 AAAAAACAATGCTGATGACACCCGAGGAGCATTTGAAGTTCCAAAGAGCTCTGGG 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 705 AAGTGAACCCCTCTTAAATGACACCGAATACAGCTGTGACCTGGAGAAACATGCCAAGGA 764

235 AAGTGTCTGTGATGACATTTGCAACCACTCAACCTGTGAGAGTACTATTTTGGC 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 765 CAAGTGTATTGACAAAGTGTGTGAACCTCAATCTTTGGAAGAAAGACTAATTGGA 824

295 CTCGAGTTTCTGATCAAAAAGATCAAGGTGTGCTGATCTCTTAAACCATTTG 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 825 CTTTGTGTTTCAGAAAGCCCTGAGCAGAAATGTGTTAGTCCGTGCTAAGAAATTAAG 884

355 AAACGATTAAGAAGCCAAAGCAGTGTGTTTAAAGTTGTGTAATTTCTTTCGGCT 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 885 AGACCAACTGAGAAACC--TTCCATGCTATTCACTTTAAATGTGAAGTTTATCTCTCT 941

415 GACCAACAACAATCTCAAGAAAGACTCAAGATGACTTGTGCGCTGAGGTGAAGCAG 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 942 GATTCCTTCAATTGACTGAAGTATACCAAGATATCTTGTGCTTCAAGCTCGGAG 1001

475 GACTTGTCTCAAGGAGGTTGACGTGTATGACACCAAGGACGCTCTTGAATTTCAAC 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1002 GACATTTGCTCGGGCGCTGCTGCTCTTTGTGATCATGCTCTCTGGGATCTTAC 1061

535 ATTGTCAATCTGAGATTGGGGATTT-----GATGAAGCCTTGAACAGAGACGACTTA 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1062 ACCCTGACAGGCTGAATCTGTGTGATGACCCAGAAACATGAGCAGATCCACCTCACT 1121

589 GCAAAAAATTAATATACATCCTCAGAAAGACGCACTAGAGGACAAATGCTGAATTTTAC 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1122 GAATTCAGTTTGGCCCTTACACTAGACTAAGAGCTGGAAGAGAGTGGCAGAGCTGCA 1181

649 CATTAACCAATTGACAAACACGACGAATCAGATTTTCAAGCTCTTGAAGTTTGG 708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1182 AAAACCCACAGGGGGCTTATCGCCAGCAGCAAGCTGATTTCCAGTTCTTGAATAATGCAAG 1241

709 CGGCTAGAGATGATGAATCCGGTTGCACCGGCCAAAGACGAGGAAGCAAGAAATC 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 1242 AGGCTTCCATGATGATGTTGACCTTACATCATGCGCAAGAGCTCAGAAAGGTGTGACATC 1301
Qy 769 AATCTGGCCCTTGGCAACACAGGGAATCTAGTGTTCAGAGGTTTCACTAAGATCATGCC 828
Db 1302 AACTGGGCGTGTGTGTCTAATGAGCTTCTCATTTACAAAGACAGACTGCCAATCAATCGT 1361
Qy 829 TTCACTGGGCGCAAGGTGTGCGGAAGCTGAGCTTCAGAGGAAGGCGTTTCTCATCAAGCTC 888
Db 1362 TTGTGCTTGGCGCAAAATCTTAAATTTCTTAAAGCGAGTAACTTCTCATTTAAAGTC 1421
Qy 889 CGGCGCATGTCCTAATAGTGTGCTACAGAGATACCTTGAATTTCTGATGCGCAAGTCGGAT 948
Db 1422 AGACCGGCGAGCTGGAACAGTTTGAAGATACCATTTGATTTCAACTGCCAAACCAACCGG 1481
Qy 949 TTCTGCAGATCTTCTGAAAAATCTGTGTGGAACATATGCTTCTTTAGACTTTTGA 1008
Db 1482 GCAGCGAAGAAAGATATGAAAGTGTGTGAGCATCATCTTCTTACAGGCTTGT--- 1538
Qy 1009 GAGCCCAACCAAGCCCAAGCCGCTCTTCTTACCGGGGGTTCATCTTTCGGTTCAGT 1068
Db 1539 TCTCCAGAGAGCCACCAAGCAAGTCTCTAAGCTTGGGGTTCCAATTTCCCTATAGT 1598
Qy 1069 GGTGGAATCTCAGAACGAG 1086
Db 1599 GCGCGACCCCAAGCAGAG 1616

RESULT 7
US-09-270-767-1995
; Sequence 1995, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1995
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1995

Query Match 4.8%; Score 165.8; DB 3; Length 576;
Best Local Similarity 61.3%; Pred. No. 4.7e-34;
Matches 285; Conservative 0; Mismatches 177; Indels 3; Gaps 1;

Qy 646 CACCATTAACCACTTGGACAAACACGAGCAATCAATTTCCAGTCCCTAGAGATTGCC 705
Db 6 CACTAAAGAGATGTGTCAATCCCGCGAGGCGGACCTTAACCTTTGAGAGCGCT 65
Qy 706 CGTGGCTAGAGATGTATGGAATCCGGTTGACCCGCGCAAGACAGGGAAGGACGAAAG 765
Db 66 CGAGGTGTGAGCTGTATGAGCAATGAAATGATCCGCGCAAGATGTGGAAGGGGTCCG 125
Qy 766 ATCAATCTGGCCGTTGCAACACGGAATTTAGTGTTCAGAGGTTTCACTAAGATCAAT 825
Db 126 CTTAATCTGGCTGTGGCCCATGAGGATCAGATCTTCCAGAAACATCAGCGGATCAAC 185
Qy 826 GCCTTCACTGGGCGCAAGGTGTGGAAGCTGAGCTTCAGAGGAAGGCGTTTCTCATCAAG 885
Db 186 ACCTTCTCGGGCTAAGATACGAAATTTCTTCAAGGCAAGGATTTCTGTCTAAA 245
Qy 886 CTCGCGCAGATGCCAATAGTGTGCTACAGATACCTTGAATTTCTGATGCGCAAGTCGG 945
Db 246 CTGCAATCCGAGGAGATATGATTTTCAAGATACCGTGAAGTTCTTCTCGAGGGTCCG 305
Qy 946 GATTTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATATGCTTCTTTAGACTTTT 1005
Db 306 AACGAGTGCMAAAATTTCTGAAAAATGCGTGCMAAATCAAGGATTTCTTCGATGCACT 365

Qy 1006 GAAGAGCCCA---ACCAAGCCCAAGCCGCTCTTAAAGCCGGGGGTTCATGATTTCCG 1062
Db 366 GCGCTCAAAATATCGCCGCGCAAAATCGGTTCTCTCGCGGGGTATGATTCGCCG 425
Qy 1063 TTCAAGTGTGCGACTCAGAAAGAGTTCGACTATATGTTAAAGAA 1107
Db 426 TATAGCGAAAAACCCAGAAAGAGATTTATCGAATGTTGTTCCGGA 470

RESULT 8
US-09-270-767-17277
; Sequence 17277, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 17277
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-17277

Query Match 4.8%; Score 165.8; DB 3; Length 576;
Best Local Similarity 61.3%; Pred. No. 4.7e-34;
Matches 285; Conservative 0; Mismatches 177; Indels 3; Gaps 1;

Qy 646 CACCATTAACCACTTGGACAAACACGAGCAATCAATTTCCAGTCCCTAGAGATTGCC 705
Db 6 CACTAAAGAGATGTGTCAATCCCGCGAGGCGGACCTTAACCTTTGAGAGCGCT 65
Qy 706 CGTGGCTAGAGATGTATGGAATCCGGTTGACCCGCGCAAGACAGGGAAGGACGAAAG 765
Db 66 CGAGGTGTGAGCTGTATGAGCAATGAAATGATCCGCGCAAGATGTGGAAGGGGTCCG 125
Qy 766 ATCAATCTGGCCGTTGCAACACGGAATTTAGTGTTCAGAGGTTTCACTAAGATCAAT 825
Db 126 CTTAATCTGGCTGTGGCCCATGAGGATCAGATCTTCCAGAAACATCAGCGGATCAAC 185
Qy 826 GCCTTCACTGGGCGCAAGGTGTGGAAGCTGAGCTTCAGAGGAAGGCGTTTCTCATCAAG 885
Db 186 ACCTTCTCGGGCTAAGATACGAAATTTCTTCAAGGCAAGGATTTCTGTCTAAA 245
Qy 886 CTCGCGCAGATGCCAATAGTGTGCTACAGATACCTTGAATTTCTGATGCGCAAGTCGG 945
Db 246 CTGCAATCCGAGGAGATATGATTTTCAAGATACCGTGAAGTTCTTCTCGAGGGTCCG 305
Qy 946 GATTTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATATGCTTCTTTAGACTTTT 1005
Db 306 AACGAGTGCMAAAATTTCTGAAAAATGCGTGCMAAATCAAGGATTTCTTCGATGCACT 365
Qy 1006 GAAGAGCCCA---ACCAAGCCCAAGCCGCTCTTAAAGCCGGGGGTTCATGATTTCCG 1062
Db 366 GCGCTCAAAATATCGCCGCGCAAAATCGGTTCTCTCGCGGGGTATGATTCGCCG 425
Qy 1063 TTCAAGTGTGCGACTCAGAAAGAGTTCGACTATATGTTAAAGAA 1107
Db 426 TATAGCGAAAAACCCAGAAAGAGATTTATCGAATGTTGTTCCGGA 470

RESULT 9
US-09-848-294-1
; Sequence 1, Application US/09848294
; Patent No. 6479640
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K.
; TITLE OF INVENTION: Isolation of A cDNA Encoding A No. 6479640el
; TITLE OF INVENTION: Protein Tyrosine Phosphatase Which Localizes to Focal

```

; TITLE OF INVENTION: Adhesions and Uses Therefor
; FILE REFERENCE: CSHL90-04FZA
; CURRENT APPLICATION NUMBER: US/09/848,294
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 09/235,251
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 08/759,536
; PRIOR FILING DATE: 1996-12-04
; PRIOR APPLICATION NUMBER: 08/107,420
; PRIOR FILING DATE: 1993-08-16
; PRIOR APPLICATION NUMBER: 07/663,579
; PRIOR FILING DATE: 1991-03-01
; PRIOR APPLICATION NUMBER: 07/494,036
; PRIOR FILING DATE: 1990-03-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3984
; TYPE: DNA
; ORGANISM: Homosapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (24) ... (2765)
; US-09-848-294-1

```

```

Query Match      4.2%; Score 144; DB 3; Length 3984;
Best Local Similarity 50.8%; Pred. No. 8.5e-28;
Matches 423; Conservative 0; Mismatches 400; Indels 9; Gaps 3;

```

```

Oy 174 CAAAATCCAGATGCTGATGACACCCAGAGGCGATTGAGTCCACAAAGAGCTCTCG 233
Db 116 CAGCATCCACTTTTATGATGGCGTGTACAGACCTTTAAAGTTACTTAACAGACACTGG 175
Oy 234 GAAGTGTCTGCTGATGACGTTTGCAACCACTCAACCTCGTGGAAAGGAGTCTATTG 293
Db 176 CCAGGTTCTTCTGATGATGTCACAAACCACTGGGTGACTGAAAGAAATATTTGG 235
Oy 294 CCTCGAGTTTCTGATCAAAAAGATCAAGG--TGCTGATCTCTTAAACCAT 350
Db 236 TTTCAGCATGATGACGATCCGTGATCTCTTATGATGGCTGGAAGCAAGCAACCAT 295
Oy 351 TGTGAAACAGATTAGAGGCCAAAGCAGTGTGTTGTAAGTTGTGTAATTTCTTCC 410
Db 296 CAGGAAGCAGTTAAAGAGGTTTCCCTGTACCTGATTTTCCAGTAAAGATTTTAT 355
Oy 411 GCTGACCAACAACAACCTCCAGAGAACTCAACAAGTACTGTTGCGCTGCAAGTGA 470
Db 356 ACCTGATCCCAACAACCTGACAGCAAGAACAAACAGGCACTGTATTTCTTCAACTGA 415
Oy 471 GCAGGACTGGCTCAAGCAGGTTGACGTGTAATGACACAGGCGAGCTCTTGAATTTC 530
Db 416 GATGATATTTGCGAAGAGAGTTAACCTGCCCTTAACTCAGCAGTGGTCTTAGCGTC 475
Oy 531 ACACATTTGCAATCTGATTTGGGATTTTGAATGAAGCTT--GACAGAGAGCACTT 587
Db 476 CTATCCCGTACATCTCATTTTGGAGACTATATTTCTTCATACATCATCCAGGCTATCT 535
Oy 588 AGCAAAAATAAATACATCTCTCAGCAAGACGCACTAGAGCAAAATGTGGAATTTCA 647
Db 536 TTCCGATGTACTTTATACCCGATCAAAATAGAGACTTTTAAACAAAATGGAATCTT 595
Oy 648 CCAT---AACCAATTTGACAAACACACAGAGATGATTTTCAAGCTCTAGAGATTTC 704
Db 596 GATAGAGCAGCAGAGGGCTTAAACATCAGAGAGAAATCTGTATATACATACATAGC 655
Oy 705 CCGTGGCTAGAGATGATGAAATCGGTTGACCCGGCCAAAGACAGGAGAGGACGAA 764
Db 656 GGGAGCCCTGACTTCTATGAGTAGAAGTCAAGTGTAGGAGATCTGCAAAATTTAGA 715
Oy 765 GATCAATCTGGCCGTGGCAACAGGGGAATCTAGTGTTCAGGGTTTCACTAAGTCA 824
Db 716 CCTAATGATTTGAAATTTGCTTCGGGGGTGTGCTGTATCCGAAATATCATTTGCAAG 775

```

```

Oy 825 TGCCTCAACTGGGCCAAGTGCAGAGCTGATGCTTCAAGAGAGCCCTTTCATCA 884
Db 776 TTTCATCTCTGGGAGAACATTTCTCAAAATTTCTTCAAAAGAAAGTTCTTATACA 835
Oy 885 GCTCCGGCCAGATGCCAATPAGTGCATACAGAGATCTTGAATTCGTATGCCAGTGC 944
Db 836 TCAGGAGACAGAAACAGGCTGATCCAGGGAACTATTGTGGCTTCAACATGCTGAATTA 895
Oy 945 GGATTTTCGAAGTCTTGTGAAATCTGTGTAATCATCATGCTTTT 996
Db 896 CCGATCTTCAAAACTTGTGAATCTGTGTGAGCACCATACGTTCTTT 947

```

```

RESULT 10
US-09-906-779-3
; Sequence 3, Application US/09906779
; Patent No. 6770466
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: Human Protein Tyrosine Phosphatase Polynucleotides, Polypeptides,
; FILE REFERENCE: PTO40Pl
; CURRENT APPLICATION NUMBER: US/09/906,779
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US01/01563
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/176,306
; PRIOR FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2872
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-906-779-3

```

```

Query Match      4.2%; Score 143.2; DB 3; Length 2872;
Best Local Similarity 50.3%; Pred. No. 1.2e-27;
Matches 490; Conservative 0; Mismatches 468; Indels 16; Gaps 5;

```

```

Oy 131 GACGAGACCGCCCCCAACACCTTCAAGAAAACCTGTGTCATCAAAATCCAGATGCTGG 190
Db 1 GGCACGAGCCCGCGCGCGCGCCAGAGCCAC--CCTCTACTGCGCGCTTCCGCTCG 59
Oy 191 ATGACACCCAGAGGCAATTTGAAGTTCACAAAGAGCTCTGGAGAAAGTGTGATG 250
Db 60 ACGGACCGAAAGTAGAGCTGACCTGCGAAACATGCAAAAGGACGAGATTTGTTGATC 119
Oy 251 CAGTTTGAACCACTCAACCTCGTGAAGGTACTATTTTGGCCTGAGTTTCTGATC 310
Db 120 AGATTGTATCAACTTGGACCTTGTGAACAGATTACTTTGGCTTCCAGTTCTGCACT 179
Oy 311 ACAAAAGATCAGGATGTGCTGATCTCTTAAACCAATGTGAAACAGATTAAGAGC 370
Db 180 CTGCCAGAGTTGCGACGTGCTGATGATGCAAAACCAATAAAAAGAGATGAATAATG 239
Oy 371 CAAGCAGTGTGTTGTAATTTTGTGTAATTTCTTCCGCTGACACACAACTCC 430
Db 240 GACCTGCTTAATCTTAACTTTGAGATTAATTAATTAATTTCTTCAAGAACCAAACTTC 299
Oy 431 AAGAGAACTCAACAGATACCTGTTGCGGCTGACAGTGAAGAGAGACTTGGCTCAAGCA 490
Db 300 GTAGAGAGTTTAAACAGATACCTGTTGTTTAACTCAAGCATGACATTTCTTCTGGA 359
Oy 491 GGTATCGTATATGACACACAGCGAGCTCTTGTATTTCAACATTTGTGCAATCTGGA 550
Db 350 AATTGAAATGCCCTTATGAAACAGCTGTGAATTAAGCTTCTGTCTACAAAGGAGAC 419
Oy 551 TTGGGATTTTGAATGAAGCTTTGACA---GAGAGCACTTGAACAAAATAAATACATAC 607
Db 420 TTGGGAGTGGAGGCTTTCAGAAACACACACAGAGCTTGTGTAGATTTGGTGTATTC 479
Oy 608 C---TCAGCAAGACGCACTAGAGCAAAATCTGTGAATTTTCACATTAACCATTTGAC 664

```

Db 480 CAATCAGACAGAGCAATGGAATTTGATCTTCCAGAGATGGAAGAGTCAGGGGAA 539
 Qy 665 AAACACCAGCAGATCAGATTTCCAGCTCTTAAGATTCCCGCTAGAGATGATG 724
 Db 540 AGAGCCCTGCCCGGGGGAATCTCTATCTGAATGAAGCGAATGGCTGGAATATG 599
 Qy 725 GAATCCGGTTGACCCGGCCAGAGACAGGGAAGCAAGATCAATCTGCGCTTGCCA 784
 Db 600 GGGTAAACATGCACTGTTTCAAGGGAAGATGGCTGTGAATTTCTCTTGAAGTAA 659
 Qy 785 ACAACGGAATTTAGTGTTCAGGTTTCACTAAGATCAATGCTTCAATGCGGCAAG 844
 Db 660 CAGCAGCATATTAATCTTTGAAGAGCTAACAAAATAGGCTATTCTTTGGCTTAAA 719
 Qy 845 TGGGAAAGCTGAGCTTCAAGAGAGAGCGCTTTCTATCAAGCTCCGCGCAGATCCA 904
 Db 720 TTAACCAAAATGATTTTAAAGAGCAAAATGACACTCGTGTGTGAGGATGATGATC 779
 Qy 905 -----GTGCGTACAGATACCTTGGAAATTCGTGATGGCAGTCCGGAATTTCTGAAGT 958
 Db 780 AGGAGCTGACAGAGACACCTTTGTTCCGTTAGACAGTGCAGAGACTTGAAC 839
 Qy 959 CCTTGTGAAAATCTGTGTGAACATCATGCTTTTGAAGTCTTTTGAAGCCCAAC 1018
 Db 840 ACTTTGGAAGTGTGAGTTGAGCACAGCATTTCTCGACTGCGGAGCCCGAGAAACA 899
 Qy 1019 CAAAGCCCAA---GCCCTCTCTTTAGCCGGGGGTCAATCTTCCGTTCACTGTGGA 1075
 Db 900 GCAAAATCCATATGATCCGACTTATATCAGCTGGGCTCGCTTCAGATTCAGTGGCGGA 959
 Qy 1076 CTCAGAGCAGGTT 1089
 Db 960 CAGAAATATCACT 973

RESULT 11
 US-09-949-016-3593
 ; Sequence 3593; Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3593
 ; LENGTH: 2156
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-3593

Query Match 4.1%; Score 142.4; DB 3; Length 2156;
 Beef Local Similarity 50.7%; Pred. No. 1.7e-27;
 Matches 422; Conservative 0; Mismatches 401; Indels 9; Gaps 3;
 Qy 174 CAAATTCAGATCTGATGATGACCCAGAGGATTTGAAGTTCCAAAGAGCTCTGG 233
 Db 110 CAGCATCCACTTTTATGATGGCGTGTACAGACCTTTAAAGTTACTAAACAGACCTGG 169
 Qy 234 GAAGTGTCTGATGATGACGTTTGAACCACTCAACCTGCTGGAAGGTGACTATTGG 293
 Db 170 CCAGGTTCTTCTGATATGATGTCACAAACCACTGGGTGTGACTGAAAAGAAATATTGG 229

Qy 294 CCTGAGTTTCTGATCAGAAAAGATCAAGG---TGGCTGATCTCTTAACCCT 350
 Db 230 TTTCAGCATGATGACGACTCCGCTGACTTCTTAAGTGGCTGGAAGCAAGACCAT 289
 Qy 351 TGTAAACAGATTGAAAGCCAAAGACGTTGTGTGAATGTTGTGTAATTTCTCC 410
 Db 290 CAGGAAGCATTTAAAGAGGTTTCCCTGTACCTTCGATTTTTCAGTAAAGATTTTAA 349
 Qy 411 GCTTACACACAACTCCAGAAAGACTCAGAGTACTTTTGGCGCTGAGGTGA 470
 Db 350 ACCTGATCCCAACACTCAGACAAACCAACAGGCACTGTATTTCTTCAACATGAA 409
 Qy 471 GCAGACTTGGCTCAAGCAGAGGTGACGTGATAGACACAGGCGAGCTCTTGAATTC 530
 Db 410 GATGATATTTGGAGAGAGGTTAACTGCTCTTAACTCAGACAGTGTTCACGCTC 469
 Qy 531 ACACATTTGCAATCTGAGATTGGGATTTTGAAGCCTT---GAGACAGAGCACTT 587
 Db 470 CTATGCCGTACATCTCATTTTGAAGATATATTTCTCATATCATTCAGGCTATCT 529
 Qy 588 AGCAAAAATTAATACATACCTCAGACAGCACTAGAGCAAAATGTGGAATTTCA 647
 Db 530 TTCCGATAGTCACTTTATACCGATCAAAATGAGACTTTTAAACAAAGTGAATCTCT 589
 Qy 648 CCAT---AACCAATTGACAAACACAGCAGATGCAATTTCCAGCTCCTAGAGATTGC 704
 Db 590 GATGAGCAGCAGAGTGGCTTAAACATCAGAGCAAGATCTGCTATATCAACTAGC 649
 Qy 705 CCGTCGCTAGAGATGATGATCCGTTTGAACCCGCGCAAGAGAGGAGGAGGAGG 764
 Db 650 GGGGACCTCGACTTCTATGAGATGAGACTGACAGTGTGAGGATCTCACAATTAGA 709
 Qy 765 GATCAATCTGCGCTTCCCAACAGGGAATTTAGTGTTCAGGTTTCACTAAGTCAA 824
 Db 710 CCTAATGATGGAATTTGCTTCGCGGGTGTGCTGTGTCCGAAAATATACATTTGCAAG 769
 Qy 825 TGCCTTCACTGGGCGCAAGTGGAGCTGAGCTTCAAGAGAGCGCTTCTCATCAA 884
 Db 770 TTTCATCTCTGGAGAACATTTCTCAAAATTTCTTCAAGAGAAAGTTCTTCTACA 829
 Qy 885 GCTCCGCGCAGATGCAATAGTGCCTACAGAGATACCTTGAATTCCTGATGGCCAGTGC 944
 Db 830 TCAGGACAGAAACAGGCTGAAATCCAGGGAACATATTTGGCGCTTCAATAGTGAATTA 889
 Qy 945 GGAATTTGCAAGTCTTCTGAAAATCTGTGTGAACATCATGCTTCTTT 966
 Db 890 CCGATCTTGCAAAAACCTGTGAAAATCTGTGTGAGCACCATACGTTCTTT 941

RESULT 12
 US-09-799-451-201
 ; Sequence 201; Application US/09799451
 ; Patent No. 6783969
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Ma, Yungting
 ; APPLICANT: Yamazaki, Victoria
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Ghosh, Reena
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. 6783969el Nucleic Acids and

TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pt_fl_genes Version 2.0
SEQ ID NO 201
LENGTH: 3398
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (174) ..(2234)
US-09-799-451-201

Query Match 3.9%; Score 133.6; DB 3; Length 3398;

Best Local Similarity 49.5%; Pred. No. 4.8e-25;
Matches 495; Conservative 0; Mismatches 484; Indels 21; Gaps 5;

185 TGCTGATGACACCCAGAGGCAATTGAAGTTCACAAAGAGCTCTGGGAAGTGTCTGC 244
115 |||||
319 TTTCTGATGTGATGATGTTAAGTGTGACCTTGCACAAAAGCCAAAGACAGAGTTGT 378
115 |||||
245 TGGATGCACTTGGCAACCACTTCACTGAGAGGTGACTATTTGGCTCGAGTTTC 304
115 |||||
379 TTGATCAGATTATGTACCACTGACCTGACTGATGAAGCAGCTATTTGGCTGAGATTTA 438
115 |||||
305 CTGATACAAAAGATCAAGGTGTGCTGATCTCTTAAACCCATTGTGAACAGATTTA 364
115 |||||
439 TGGATTACAGACAGAGTAGACATTTGGTGTGATGTACAAAAGATCAAAAAGCAGTAA 498
115 |||||
365 GAAGGCCAAAGACGTTGTGTTAAGTTGTGTTAATTTCTTCCGCTGACCAACAC 424
115 |||||
499 AATTGTGTTACCTTATTTGTCTGATCTTCAAGTTAAGTTTATCTCAGAACCAATA 558
115 |||||
425 AACTCCAAAGAACTCACAAGGTACCTGTGCGCTGCAAGGTGAAGCAGAGCTTGGCTC 484
115 |||||
559 ACCTTGATGAGAGCTAACCCGGTATTTATTTGTTCTCAGTTAAACAAAGATATTCTCA 618
115 |||||
485 AAGGCGAGTTGACGTGTATGACACAGAGGAGCTCTGATTTTACACATTGGCAAT 544
115 |||||
619 GTGGAAATTTAAGCTGTCTCTTGTATACAGAGTGAATTTGGAGTTTAAATCTGCAAG 678
115 |||||
545 CTGAGATTTGGGATTTTGTGAAGCCTTGGACAG--AGAGCACTTAGCAAAAATAAT 601
115 |||||
679 CTGAACTTGTGATGATGATCTTGTGAGCATTAGTCTGAACTTGTCTCAGAGTTCAAGT 738
115 |||||
602 ACATACCC--TGACAAAGCGCACTAAGAGGACAAAATCGTGAATTTCAACATAACACA 658
115 |||||
739 TCGTGCTATTGACAGTGAAGAGATGGAATCGGCTATTTTGAAGAAATGGAAGATATCA 798
115 |||||
659 TTGGAACAACACAGAGAGATCAGATTTCCAGCTCTTAAGATGTCGCGCTGAGAGAA 718
115 |||||
799 GAGGTCAAAACACAGACAGCTGAAAACCAATTATCTGAATTAAGCAAAATGGCTAGAAA 858
115 |||||
719 TGTATGGAATCCGGTTGACACCGGCCAAGACAGGAAAGCAAGATCAATCTGGCCG 778
115 |||||
859 TGTATGGGTTGATATGATGCTATGCTAAGCTAAGATGGAATGATCTAATGTTGGAGC 918
115 |||||
779 TTGCAACACGAGGAAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAATGCG 838
115 |||||
919 TAAACACCAAGAGAGCTGTGTTTGAAGAGATACCAAAATTTGGCTTATTTTGGC 978
115 |||||
839 CCAAGGTGGGAAGCTGAGCTTCAAGAGAGAGCGCTTCTCAATCAAGCTCCGCGCAGAG 898
115 |||||
979 CCAAGATTAACAGATTTGATTTTAAAGAAATAATTAACCTTGTGTGTGTAAGAGATG 1038
115 |||||
899 CCAAT-----AGTGGTACCAAGATACCTTGAATCTCTGATGCGCAGTCCGGAATTTCT 952
115 |||||
1039 ATGATCAGGCGCAAGAAAGAGAAACATATTTGTCTTGAATCAGATCATCCAAAAGCAT 1098
115 |||||
953 GCAAGTCTTTTGGAAATCTGTGTTGAACATCATGCTCTTTTGAATCTTTTGAAGAGC 1012
115 |||||

1099 GCAAACTTTATGGAATGTGCTGTGAGCATCATGCTTTCTCCGCTTCAGAGCCCG 1158
115 |||||
1013 CCAACCAAGGCCCAAGC---CCGCTCTTTAGCCGGGGGATCATTTGGTTCAAGT 1069
115 |||||
1159 TCCAAAGGTTTCTCATGATGATGATTTATTCGATGAGATCAGATTTAATATGATG 1218
115 |||||
1070 GTGGAAGTCAAGAGAGGTTTCTGACTATGTTTAAAGAGAGAGATTAAGAGTCACT 1129
115 |||||
1219 GGAAGACAGATATGAGACCAACAAACCAATTAAGAGAGATCA-----CATCT 1272
115 |||||
1130 TTGAAGAGAGACAGCAAGATTCTTCTATCCGAGCTT 1169
115 |||||
1273 TTGAAGAGAGGCCCGCAAGCATATTTCTAGCAGAACTT 1312
115 |||||

RESULT 13

US-09-270-767-1542
Sequence 1542, Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1542

LENGTH: 449
TYPE: DNA

ORGANISM: Drosophila melanogaster
US-09-270-767-1542

Query Match 3.1%; Score 106.4; DB 3; Length 449;

Best Local Similarity 65.7%; Pred. No. 3.4e-18;
Matches 155; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

2761 CCGCACCCGCGCACACATGATGATGCTGTGTCGACCCGCAACAGCAGCTTCATG 2820
115 |||||
199 CCATCCCAAGAAATACCGGCTGTGACGTCGTCGATTCGATTCGAGGCGCACCGTAGACTG 258
115 |||||
2821 GTGACTTCAGATGACAGTGAAGATCAATGTTGTGAAACCTGCTGAGAAATTCAA 2880
115 |||||
259 GCGCATCACCTGATATCCGCGCAGACCAAGTTGTCGCGTATATCTGTGAAATGCAAG 318
115 |||||
2881 AACAGCAAGGTTGCAAGCTGTGAGGTTGTTCAAACTTGTGCTGCTGCTGCTTCTAC 2940
115 |||||
319 AATACCTCCGCTGCAAAAGCTGTGGGTGTGATGCTCTTGTGTGATCTTTAC 378
115 |||||
2941 AATACACACAGCAATATCTCCCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 2996
115 |||||
379 AAAACCTACAGAGAGAGATTCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 434
115 |||||

RESULT 14

US-09-270-767-16824
Sequence 16824, Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16824

LENGTH: 449
TYPE: DNA

ORGANISM: Drosophila melanogaster
US-09-270-767-16824

Query Match 3.1%; Score 106.4; DB 3; Length 449;

Best Local Similarity 65.7%; Pred. No. 3,4e-18;
Matches 155; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 2761 CCGCACC GCGGCAACACATGTCACGTGTCTGGCACC GCAACACAGCGTCTCCATG 2820
DB 199 CCATCCCGAAGCAATAGCGGCTCTGTCACGTCTGTGGCATCGTGCGCCACCGTAGGACTG 258
QY 2821 GTGGACTTCAGCATCCGAGTGGAGAAATCATGTTGTTCTGAAAACTGCTGAGAGAAATTCAAA 2880
DB 259 GCGGATTCACCTGATATCCGCCGAGCACAGTGTCCGGTTATCTGCTGCGAAAGTTCAG 318
QY 2881 AACAGCAACGGGTGGCAGAACTGTGGGTGTGTTCAAACTTCTGCTGTTCTTCTAC 2940
DB 319 AATAGCTCCGGCTGGCAAAAGCTCTGGGTGTGTTCAAGTCTTCTGTAATTTTAC 378
QY 2941 AATTCACACCAGACAAATCATCCCTTGGCAGCCCTGCTGCTGCGCTACTGCGT 2996
DB 379 AAAAGCTACGAGACGAGTTCGACATGGCTAGTCTGCGCTATGGGTACACGGT 434

RESULT 15

US-09-621-976-15728
; Sequence 15728, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15728
; LENGTH: 184
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15728

Query Match 3.1%; Score 105.2; DB 3; Length 184;
Best Local Similarity 99.1%; Pred. No. 4.5e-18;
Matches 105; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1740 GTGGTTTCAGAGCACAGTAGCAAGAGAGCGCATGCGGAAAGCACTGAAAAAGTCTCAT 1799
DB 79 GTGGTTTCAGAGCACAGTAGCAAGAGAGCGCATGCGGAAAGCACTGAAAAAGTCTCAT 138
QY 1800 ATTCCGGAATTTGAACCTTGGCAAAATTTCACTAATTTTCTC 1845
DB 139 ATTCCGGAATTTGADCTTTGCAAAATTTCACTAATTTTCTC 184

Search completed: May 22, 2006, 09:18:19
Job time : 597.56 secs

This Page Blank (uspto)

181 CAGATGCTGATGACACCCGAGAGCATTTGAAGTTCCAAAGAAGCTCTGGGAAGTG 240
181 CAGATGCTGATGATGACCCGAGAGGCAATTTGAAGTTCCAAAGAAGCTCTGGGAAGTG 240
241 CTGCTGATGATGAGTTTGCAACCACTCAACTCTGTGGAAGGTGACTATTTTGGCTCGAG 300
241 CTGCTGATGATGAGTTTGCAACCACTCAACTCTGTGGAAGGTGACTATTTTGGCTCGAG 300
301 TTTCCTGATCAGAAAAGATCAAGGTGAGTGGTGGATCTCTAAACCCCATTTGGAACAG 360
301 TTTCCTGATCAGAAAAGATCAAGGTGAGTGGTGGATCTCTAAACCCCATTTGGAACAG 360
361 ATTAGAAGCCAAAGACGTTGTTGTTAAAGTTGTGTAAGAAATCTTTCCGCTGACAC 420
361 ATTAGAAGCCAAAGACGTTGTTGTTAAAGTTGTGTAAGAAATCTTTCCGCTGACAC 420
421 ACACAATCTCAAGAAAGATCTCAAGAGTACCTGTTGCGCTGCAAGGTGAAGACATTTG 480
421 ACACAATCTCAAGAAAGATCTCAAGAGTACCTGTTGCGCTGCAAGGTGAAGACATTTG 480
481 GCTCAAGGCGATGACGTGTAATGACACAGCGGACGCTCTTGATTTTCAACAATGTG 540
481 GCTCAAGGCGATGACGTGTAATGACACAGCGGACGCTCTTGATTTTCAACAATGTG 540
541 CAATCTGAGATTTGGGATTTTGAATGAAGCTTGAACAGAGACATTTAGCAAAAATAA 600
541 CAATCTGAGATTTGGGATTTTGAATGAAGCTTGAACAGAGACATTTAGCAAAAATAA 600
601 TACATACCTCAGCAAGACGACCTAGAGGACAAATCTGTGAATTTCAACATTAACCAAT 660
601 TACATACCTCAGCAAGACGACCTAGAGGACAAATCTGTGAATTTCAACATTAACCAAT 660
661 GGACAAACACGAGCAATCAGATTTCCAGCTCTAGAGATTTGCCGTGAGTGAAGATG 720
661 GGACAAACACGAGCAATCAGATTTCCAGCTCTAGAGATTTGCCGTGAGTGAAGATG 720
721 TATGGAATCCGGTTGACCCGAGCAAGGACAGGAAAGGACAGAAAGTCAATCTGGCCGTT 780
721 TATGGAATCCGGTTGACCCGAGCAAGGACAGGAAAGGACAGAAAGTCAATCTGGCCGTT 780
781 GCCAACACGAGGAATCTAGTGTGTTTCAAGGTTTCAAGATCAATGCTTCAACTGGGCC 840
781 GCCAACACGAGGAATCTAGTGTGTTTCAAGGTTTCAAGATCAATGCTTCAACTGGGCC 840
841 GGCACACGAGGAATCTAGTGTGTTTCAAGGTTTCAAGATCAATGCTTCAACTGGGCC 840
841 GGCACACGAGGAATCTAGTGTGTTTCAAGGTTTCAAGATCAATGCTTCAACTGGGCC 840
841 AAGGTGCGGAAGCTGAGCTTCAAGAGAGAGCGCTTTCATCAAGTCCGAGCAGATGTC 900
841 AAGGTGCGGAAGCTGAGCTTCAAGAGAGAGCGCTTTCATCAAGTCCGAGCAGATGTC 900
901 AATAGTGCATACAGAGATACCTTGAATTCCTGATGCGCAGTCGGGATTTCTGCAAGTCC 960
901 AATAGTGCATACAGAGATACCTTGAATTCCTGATGCGCAGTCGGGATTTCTGCAAGTCC 960
961 TTCTGGAATAATCTGTGTTGAACATCATGCTTCTTTAGACTTTTGAAGAGCCCAACCA 1020
961 TTCTGGAATAATCTGTGTTGAACATCATGCTTCTTTAGACTTTTGAAGAGCCCAACCA 1020
1021 AAGCCCAAGCCGCTCTTTAGCCGAGGATCATCATTTGCGTTCAAGTGTGAGCTCAG 1080
1021 AAGCCCAAGCCGCTCTTTAGCCGAGGATCATCATTTGCGTTCAAGTGTGAGCTCAG 1080
1081 AAGCAGGTTCTGACTATGTAAAGAGAGACATTAAGAGGTGCAATTTGAAGAGAG 1140
1081 AAGCAGGTTCTGACTATGTAAAGAGAGACATTAAGAGGTGCAATTTGAAGAGAG 1140
1141 CACAGCAAGATTCATTTCTATCCGAGGCTTTGCTTCAAGCTTCAAGAACTGAATTCGGA 1200
1141 CACAGCAAGATTCATTTCTATCCGAGGCTTTGCTTCAAGCTTCAAGAACTGAATTCGGA 1200
1201 GTGCTGAGAGCTCTCAGAGAGACACAGCTTCACTTTGAGAGAGGTCCGAACTCTCA 1260
1201 GTGCTGAGAGCTCTCAGAGAGACACAGCTTCACTTTGAGAGAGGTCCGAACTCTCA 1260
1261 GGGGAGCAGAGCTGCGCGGAGAGAAAGAAAGGATTTCCGCGGAGAGCCGAGGCTCG 1320

1261 GGGGAGCAGAGCTGCGCGGAGAGAAAGGAAAGGATTTCCGCGGAGAGCCGAGGCTCG 1320
1321 CACCCGAGCCCTGCGCCGAGAGAAAGCCCGCGGTTAAACAGACAGCCGAGACCCGCG 1380
1321 CACCCGAGCCCTGCGCCGAGAGAAAGCCCGCGGTTAAACAGACAGCCGAGACCCGCG 1380
1381 TCGGCGCCCAAGAGAGAGAGAGGTCTGTTAAGATTAAGAACCCAGAGAGTAACT 1440
1381 TCGGCGCCCAAGAGAGAGAGAGGTCTGTTAAGATTAAGAACCCAGAGAGTAACT 1440
1441 CAGCCCGGAGCCAGAGACAGAGCTCCCTGACCTGCACTGCTCCTTTCCGAGCTGCT 1500
1441 CAGCCCGGAGCCAGAGACAGAGCTCCCTGACCTGCACTGCTCCTTTCCGAGCTGCT 1500
1501 GTGAATCTGAGAGGAGAGGTGCGCTTCCAGACGATCTTGTCTCCCACTGAGCCCC 1560
1501 GTGAATCTGAGAGGAGAGGTGCGCTTCCAGACGATCTTGTCTCTCCCACTGAGCCCC 1560
1561 GACACCAAGCAGGCTCTCCCTGATCAGCCGCTGCTGAATGAACAGGCTGCGCCCGG 1620
1561 GACACCAAGCAGGCTCTCTCTGATCAGCCGCTGCTGATGAACAGGCTGCGCCCGG 1620
1621 ACGAGATGAGATGAGAGGCGCGAGAGAGATTTCCCACTGATTAAGCTTCTCAT 1680
1621 ACGAGATGAGATGAGAGGCGCGAGAGAGATTTCCCACTGATTAAGCTTCTCAT 1680
1681 GCTAAGAGAGTCTACACCGAGAGGACATATCTGAAGATCTGAGATTAATCACTTCG 1740
1681 GCTAAGAGAGTCTACACCGAGAGGACATATCTGAAGATCTGAGATTAATCACTTCG 1740
1741 TGTGTTTCAAGACACAGTGAAGAGAGAGCCCATGCGGAAAGCACTGAATTAAGTCTCAT 1800
1741 TGTGTTTCAAGACACAGTGAAGAGAGAGCCCATGCGGAAAGCACTGAATTAAGTCTCAT 1800
1801 TTCCGAATTTTGAACCTTTGACAAATTTCTAATTAATTTCTCAAGAAATTAAGCA 1860
1801 TTCCGAATTTTGAACCTTTGACAAATTTCTAATTAATTTCTCAAGAAATTAAGCA 1860
1861 CGACTTGGCCCTGTGGAGAGGCGCTCAAAATGCCAAATCAGAGATTAACAAAGATCGGC 1920
1861 CGACTTGGCCCTGTGGAGAGGCGCTCAAAATGCCAAATCAGAGATTAACAAAGATCGGC 1920
1921 GATGTCATGCTGAAGAAATTCAGGGCATGAAGCACTGAGGCTCACTGTGAAGAGC 1980
1921 GATGTCATGCTGAAGAAATTCAGGGCATGAAGCACTGAGGCTCACTGTGAAGAGC 1980
1981 AAGCAGGCTTTGAGAGCCCTGAGAGATGAATCAAGAGCTTCCGAGCTGAGAACTTC 2040
1981 AAGCAGGCTTTGAGAGCCCTGAGAGATGAATCAAGAGCTTCCGAGCTGAGAACTTC 2040
2041 TGCGAGACTTTGAGCTGACAGAGGTGTTAACCCTACCGCTCAACCTTCTCTGGGG 2100
2041 TGCGAGACTTTGAGCTGACAGAGGTGTTAACCCTACCGCTCAACCTTCTCTGGGG 2100
2101 CCAGTGCACCGGCTCATAGCACTACAGAGGCTCTGAGAGCGCTGTGCAACACCAACCG 2160
2101 CCAGTGCACCGGCTCATAGCACTACAGAGGCTCTGAGAGCGCTGTGCAACACCAACCG 2160
2161 CCGAGCCACGCGGCTTCAAGGAGCTGCGAGGCGCTTTGGCAGAGATCAAGAGATGTTG 2220
2161 CCGAGCCACGCGGCTTCAAGGAGCTGCGAGGCGCTTTGGCAGAGATCAAGAGATGTTG 2220
2221 GCAAGCTCCACCGGATGATGATCAAGATGAATGAATTTCAAGAGCTCAAGAGATCAAG 2280
2221 GCAAGCTCCACCGGATGATGATCAAGATGAATGAATTTCAAGAGCTCAAGAGATCAAG 2280
2281 AAAGATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
2281 AAAGATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
2341 AGCTTCAAGAGCTTCTCGGGAGAGGAGCTTCAAGAGGCTGATTTTCTGTTCAAGAG 2400

Db 2341 AGCCTCAGACAGCTCTCGGGGAGAGGGGCTCAGAGGCACTTCTTCTCTTCAACGAC 2400
Qy 2401 GTCTCTGATATACAGAGCCGGGGGCTGACGGCTCCATATAGTTTAAAGTCCAGGGCAG 2460
Db 2401 GTCTCTGATATACAGAGCCGGGGGCTGACGGCTCCATATAGTTTAAAGTCCAGGGCAG 2460
Qy 2461 CTCGCGCTCATATGAGATGAGATTGAGAGGAGGAGAGAGAGGGGGGGGCGCCACTGC 2520
Db 2461 CTCGCGCTCATATGAGATGAGATTGAGAGGAGGAGAGAGAGGGGGGGGCGCCACTGC 2520
Qy 2521 CTGACCTCCGGGGGCGAGCGGAGCTCATCATCTGTGGCCGCACTTCTCGGTCCAGATG 2580
Db 2521 CTGACCTCCGGGGGCGAGCGGAGCTCATCATCTGTGGCCGCACTTCTCGGTCCAGATG 2580
Qy 2581 GAGAAATGGGTTGAGAGCATCCAGATGGCCATTGACTGGCGGAGAGAGACAGACGCC 2640
Db 2581 GAGAAATGGGTTGAGAGCATCCAGATGGCCATTGACTGGCGGAGAGAGAGACAGACGCC 2640
Qy 2641 GCGCCCTGAGTTCCTGGCCAGACGCCCCCTGACAAAGTCCGCTGATGAGAGCAGCGG 2700
Db 2641 GCGCCCTGAGTTCCTGGCCAGACGCCCCCTGACAAAGTCCGCTGATGAGAGCAGCGG 2700
Qy 2701 GCTGACAGAGATCAGAGATGACCTGAGCGCTCGCGACATCGCTGAGAGCGCCAGGC 2760
Db 2701 GCTGACAGAGATCAGAGATGACCTGAGCGCTCGCGACATCGCTGAGAGCGCCAGGC 2760
Qy 2761 CCGCACCGGGGCAACAAATGGTGCACGTGTGCTGGACCGCAACACAGCGTCTCAATG 2820
Db 2761 CCGCACCGGGGCAACAAATGGTGCACGTGTGCTGGACCGCAACACAGCGTCTCAATG 2820
Qy 2821 GTGGAATTGAGATCGCAGATGAGAGATGAGTGTCTGAGAACTGCTGAGAGAAATTCAA 2880
Db 2821 GTGGAATTGAGATCGCAGATGAGAGATGAGTGTCTGAGAACTGCTGAGAGAAATTCAA 2880
Qy 2881 AACAGCAAGGGTGGAGAGAGCTGTGGTGTGTTCACAACTTCTGCTTCTTCTAC 2940
Db 2881 AACAGCAAGGGTGGAGAGAGCTGTGGTGTGTTCACAACTTCTGCTTCTTCTAC 2940
Qy 2941 AATATCAGACAGAGACATATCCCCCTTGGCAGCGCTGCTGTGGCTATCTGCTAC 3000
Db 2941 AATATCAGACAGAGACATATCCCCCTTGGCAGCGCTGCTGTGGCTATCTGCTAC 3000
Qy 3001 ATCCCTCTGAGTCCGAGAACATCCAGAAAGCTAGTGTCAAGCTGCATTCAAGTCC 3060
Db 3001 ATCCCTCTGAGTCCGAGAACATCCAGAAAGCTAGTGTCAAGCTGCATTCAAGTCC 3060
Qy 3061 CAGCTCTACTACTTCAAGGCGGAGAAAGCAGTACACGTTGCAAAAGTGTGATGATC 3120
Db 3061 CAGCTCTACTACTTCAAGGCGGAGAAAGCAGTACACGTTGCAAAAGTGTGATGATC 3120
Qy 3121 CCGAGTGCACACAGCTCTGCTGCGACCCCAAGTGTGAGCCAAAGAGTCTCTTG 3180
Db 3121 CCGAGTGCACACAGCTCTGCTGCGACCCCAAGTGTGAGCCAAAGAGTCTCTTG 3180
Qy 3181 TATTGATGGCCGGACACCTCGTTCCGAGTGGCTGCTTCTGGAAGAGCTTCTCTT 3240
Db 3181 TATTGATGGCCGGACACCTCGTTCCGAGTGGCTGCTTCTGGAAGAGCTTCTCTT 3240
Qy 3241 CTTCTGTATTAATGAGAGCTGTGTAATTAACACTGTCTGAAATCAAAAATCATGGCTT 3300
Db 3241 CTTCTGTATTAATGAGAGCTGTGTAATTAACACTGTCTGAAATCAAAAATCATGGCTT 3300
Qy 3301 CCGAGAGCTCTCTCTGTCCACAGCGGGTTTTTAAACCCGAGCTCTCAGCGTTTGA 3360
Db 3301 CCGAGAGCTCTCTCTGTCCACAGCGGGTTTTTAAACCCGAGCTCTCAGCGTTTGA 3360
Qy 3361 TGAACAGCGCTCCACCTCAGTCTGGATCCGCTGGGGGGGCTGTCTTTAGTATG 3420
Db 3361 TGAACAGCGCTCCACCTCAGTCTGGATCCGCTGGGGGGGCTGTCTTTAGTATG 3420
Qy 3421 CCAGTATTAACATTTGCTATT 3442
Db 3421 CCAGTATTAACATTTGCTATT 3442

RESULT 2
US-10-921-707-18
; Sequence 18, Application US/10921707
; Publication No. US20050003447A1
GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BANDMAN, Olga
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0594 PCT
; CURRENT APPLICATION NUMBER: US/10/921,707
; CURRENT FILING DATE: 2004-08-19
; PRIOR APPLICATION NUMBER: US/09/786,797
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: 09/156,470; unassigned: 60/131,321
; PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 18
; LENGTH: 4687
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1320252
US-10-921-707-18
Query Match 99.9%; Score 3437.2; DB 9; Length 4687;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3439; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CCGCGCAGCGCGCGCGCTGTGAGATATTTCTTAAGCGCTTTCATCATGAGAGAAATA 60
Db 42 CCGCGCAGCGCGCGCGCTGTGAGATATTTCTTAAGCGCTTTCATCATGAGAGAAATA 101
Qy 61 GAGCAGAGCGCGAGCCCGAGATCAGCATGAGGCGCCCGAGAAATTCGGGATCAGTACC 120
Db 102 GAGCAGAGCGCGAGCCCGAGATCAGCATGAGGCGCCCGAGAAATTCGGGATCAGTACC 161
Qy 121 TTGGAACGTGGAACAGAGCGCGCCCGCAACACTTCAGAGAAAATCGGTGTCTCAAAATC 180
Db 162 TTGGAACGTGGAACAGAGCGCGCCCGCAACACTTCAGAGAAAATCGGTGTCTCAAAATC 221
Qy 181 CAGATGCTGATGACACCCAGAGAGCATTTGAAGTTTCCCAAAAGAGCTCTGGAAGGTG 240
Db 222 CAGATGCTGATGACACCCAGAGAGCATTTGAAGTTTCCCAAAAGAGCTCTGGAAGGTG 281
Qy 241 CTGTGTGATGACATTTGCAACCACTCAACTGTGGAAGGTGACTATTTTGGCTCGAG 300
Db 282 CTGTGTGATGACATTTGCAACCACTCAACTGTGGAAGGTGACTATTTTGGCTCGAG 341
Qy 301 TTTCTGTGATCAAAAAGATCAACGTTGTGCTGTGATCTTCTTAAACCATTTGTGAACAG 360
Db 342 TTTCTGTGATCAAAAAGATCAACGTTGTGCTGTGATCTTCTTAAACCATTTGTGAACAG 401
Qy 361 ATTGAAGGCGAAAGCAGCTGTGTGTAAGTTTGTGTAATTTCTTCCGCTGACAC 420
Db 402 ATTGAAGGCGAAAGCAGCTGTGTGTAAGTTTGTGTAATTTCTTCCGCTGACAC 461
Qy 421 ACACAACTCCAAAGAAATCAACAAAGTACTGTTTGGCTGACAGTGAAGCAGACTTG 480
Db 462 ACACAACTCCAAAGAAATCAACAAAGTACTGTTTGGCTGACAGTGAAGCAGACTTG 521

OY	481	ICTCAAGCAGAGTTGACGTCTGAATGACAACAGCGCACGCTCTCTTGATTTCACATTTGYG	540
Db	522	GCTCAAGGCAGAGTTGACGTCTGAATGACAACAGCGCACGCTCTCTTGATTTCACATTTGYG	581
OY	541	CAATCTGAGATTGGGGATTTTGTATGAAGCCCTTGGACAGAGAGCACTTAGCAAAATAA	600
Db	582	CAATCTGAGATTGGGGATTTTGTATGAAGCCCTTGGACAGAGAGCACTTAGCAAAATAA	641
OY	601	TACATPACTTCAGCAAGACGCACTAGAGGACAAAATCGTGAATTTCAACATPACCACTT	660
Db	642	TACATPACTTCAGCAAGACGCACTAGAGGACAAAATCGTGAATTTCAACATPACCACTT	701
OY	661	GGACAAACACAGAGCAATCAATTTTCCAGTCTCTAGAGATTGGCCCGCGGCTAGAGATG	720
Db	702	GGACAAACACAGAGCAATCAATTTTCCAGTCTCTAGAGATTGGCCCGCGGCTAGAGATG	761
OY	721	TATGGAATCCGGTTGCAACCCGGCCAGAAGAGGAGCAAGAGATCAATCTGAGCGTT	780
Db	762	TATGGAATCCGGTTGCAACCCGGCCAGAAGAGGAGCAAGAGATCAATCTGAGCGTT	821
OY	781	GCCAACAACGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCCCTTCAACTGAGCC	840
Db	822	GCCAACAACGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCCCTTCAACTGAGCC	881
OY	841	AAGGTGCGGAAGCTGAGAGTTCAAGAGGAAGCGCTTCTCATCAAGCTCCGGCCAGATGCC	900
Db	882	AAGGTGCGGAAGCTGAGAGTTCAAGAGGAAGCGCTTCTCATCAAGCTCCGGCCAGATGCC	941
OY	901	AATATGTGCTACACAGATACCCTTGAATTCCTGATGGCCAGTCCGGGAATTTCTGCAAGTCC	960
Db	942	AATATGTGCTACACAGATACCCTTGAATTCCTGATGGCCAGTCCGGGAATTTCTGCAAGTCC	1001
OY	961	TTCTGAAAAATCTGTGTGAAACATCATGCTCTTCTTGAACCTTTTGAAGAGCCCAAACA	1020
Db	1002	TTCTGAAAAATCTGTGTGAAACATCATGCTCTTCTTGAACCTTTTGAAGAGCCCAAACA	1061
OY	1021	AAGCCCAAGCCCGCTCTTTTAAGCCGGGGGTCATCATTTCCGTTCAAGTGTGCGACTCAG	1080
Db	1062	AAGCCCAAGCCCGCTCTTTTAAGCCGGGGGTCATCATTTCCGTTCAAGTGTGCGACTCAG	1121
OY	1081	AAGCAGGTTCTCGACTATGTTAAAGAGAGAGACATAAGAAAGTGACTTTGAAGAGAG	1140
Db	1122	AAGCAGGTTCTCGACTATGTTAAAGAGAGAGACATAAGAAAGTGACTTTGAAGAGAG	1181
OY	1141	CACAGCAAGATTCTATCTATCCGAGCCCTTGCTTCAAGCCTTACAGAACATGAAATTCGAAA	1200
Db	1182	CACAGCAAGATTCTATCTATCCGAGCCCTTGCTTCAAGCCTTACAGAACATGAAATTCGAAA	1241
OY	1201	GTCGTTGAGCAGTCTCAGCAGAGCACCAAGCCTTACATTTGAGAAAGGTGCCAATCTTCCA	1260
Db	1242	GTCGTTGAGCAGTCTCAGCAGAGCACCAAGCCTTACATTTGAGAAAGGTGCCAATCTTCCA	1301
OY	1261	GGGGGGCAGABACTCCGGCGAGAGAAAGAACCGAAGATTTCCGCGGGGAGACCGGGGGTGC	1320
Db	1302	GGGGGGCAGABACTCCGGCGAGAGAAAGAACCGAAGATTTCCGCGGGGAGACCGGGGGTGC	1361
OY	1321	CACCCGAGCCCTTGCGCCGAGAGAGAACCCCGCGGGTTAAACAAGACAGGCGGACCGCGCC	1380
Db	1362	CACCCGAGCCCTTGCGCCGAGAGAGAACCCCGCGGGTTAAACAAGACAGGCGGACCGCGCC	1421
OY	1381	TCGGCGCCCAACGAGAGAAAGAGAGAGGTGCTTAAGATAGAGACCCAGAGAGTTAAACT	1440
Db	1422	TCGGCGCCCAACGAGAGAAAGAGAGAGGTGCTTAAGATAGAGACCCAGAGAGTTAAACT	1481
OY	1441	CAGCCCCCGCAGCCAAAGCAGAGCTTCCCTGACTGGCACTCTTACCTTTCCGACTGTCT	1500
Db	1482	CAGCCCCCGCAGCCAAAGCAGAGCTTCCCTGACTGGCACTCTTACCTTTCCGACTGTCT	1541
OY	1501	GTGAACCTGCAAGGGGGAGTGGCCCTTGCAACGTGACTTGTCTCCCAACCTTAGAGCCC	1560
Db	1542	GTGAACCTGCAAGGGGGAGTGGCCCTTGCAACGTGACTTGTCTCCCAACCTTAGAGCCC	1601
OY	1561	GACACCAAGAGGCTCTCCCTTGAATGACCCGCTGCTGAATGACAGAGGCTGCCCCCG	1620

Dp	1602	GACCAAGCAGGACCTTCTCCCTTGATAGCCCGCTGTAATGACGAGGCTGCCCCGG	1661
Qy	1621	ACGCAACATGAGATGATGAGGCGCGAGGAGAAGATTCCCACTGATAAAGCTACTTCATA	1680
Dp	1662	ACGGACGATGAGGATGAGGGCCGGAGAGAGATTTCCCACTGATTAAGGCTACTTCATA	1721
Qy	1681	GCTAAGGAATGCTCTACCAACCGAGCGAACTATCTGTAAGGATCTCCGAAGTTATCACTTCG	1740
Dp	1722	GCTAAGGAAGTGTCTTACCCACCGAGCGAACTATCTGTAAGGATCTCCGAAGTTATCACTTCG	1781
Qy	1741	TGTTTTGAGACACAGTGAAGCAAAAGAGAGCGCCATGCGGAAAGCATGAAAAGTCTCAT	1800
Dp	1782	TGTTTTGAGACACAGTGAAGCAAAAGAGAGCGCCATGCGGAAAGCATGAAAAGTCTCAT	1841
Qy	1801	TTCCCGAATTTTGAAACCTTTGCACAAATTTCACTAATTTCTCAAGAAATTTGACAA	1860
Dp	1842	TTCCCGAATTTTGAAACCTTTGCACAAATTTCACTAATTTCTCAAGAAATTTGACAA	1901
Qy	1861	CGACTTCGCTGTGTGGAAAGCGCGCTCAATATGCCCAATCAAGATTTCAAAAGATGGCG	1920
Dp	1902	CGACTTCGCTGTGTGGAAAGCGCGCTCAATATGCCCAATCAAGATTTCAAAAGATGGCG	1961
Qy	1921	GATGTCATGTGTAAGAACTTTCAGGGCATAGAGCACTGCGCGCTCACTGTGGAACAC	1980
Dp	1962	GATGTCATGTGTAAGAACTTTCAGGGCATAGAGCACTGCGCGCTCACTGTGTGAACAC	2021
Qy	1981	AGCAGGCGCTTGTGAGGCGCTGTGGAATGGAATCAAGACTCCGCGCGCTGTGGAATCTT	2040
Dp	2022	AGCAGGCGCTTGTGAGGCGCTGTGGAATGGAATCAAGACTCCGCGCGCTGTGGAATCTT	2081
Qy	2041	TGCAGAGACTTTGAGCTGCAAGAAAGTGTGTTTACTTACCGCTCAACCTTCTCTGCGG	2100
Dp	2082	TGCAGAGACTTTGAGCTGCAAGAAAGTGTGTTTACTTACCGCTCAACCTTCTCTGCGG	2141
Qy	2101	CCACTGCAACGGGCTCATAGCACTTCAAGCAGGCTCTGAGGGCGGTGCAAAACACACCG	2160
Dp	2142	CCACTGCAACGGGCTCATAGCACTTCAAGCAGGCTCTGAGGGCGGTGCAAAACACACCG	2201
Qy	2161	CCGAGCAACGCGCACTTCAAGGACCTGCGCGAGCGCTTTGGCAGAGATCAAGAGATGCTG	2220
Dp	2202	CCGAGCAACGCGCACTTCAAGGACCTGCGCGAGCGCTTTGGCAGAGATCAAGAGATGCTG	2261
Qy	2221	GCACAGCTTCCACGGTACGATGATCAAGATGAGAGATTTCCAGAAAGCTGCACGAACTCAAG	2280
Dp	2262	GCACAGCTTCCACGGTACGATGATCAAGATGAGAGATTTCCAGAAAGCTGCACGAACTCAAG	2321
Qy	2281	AAAGATTTGATTTGGCAATTTGCAATCTTGTGTGTTCCGGGAAGGAGTTCAATCCGCTGGCG	2340
Dp	2322	AAAGATTTGATTTGGCAATTTGCAATCTTGTGTGTTCCGGGAAGGAGTTCAATCCGCTGGCG	2381
Qy	2341	AGCCTCAAGCAAGCTCTCGGGGAAGGGGCTCAGACAGGCACTGTTCTTCTGTTCAAGCAG	2400
Dp	2382	AGCCTCAAGCAAGCTCTCGGGGAAGGGGCTCAGACAGGCACTGTTCTTCTGTTCAAGCAG	2441
Qy	2401	GTCCTGTCTATACAGCAGCCGAGGGCTGACGCGCTCCAACTAGTTTAAAGTTCACGAGCAG	2460
Dp	2442	GTCCTGTCTATACAGCAGCCGAGGGCTGACGCGCTCCAACTAGTTTAAAGTTCACGAGCAG	2501
Qy	2461	CTCCCCGCTCTATATGCAATGATTTGAGAGAGCGAAGACGAGTGGGGGTGCCCCACTGC	2520
Dp	2502	CTCCCCGCTCTATATGCAATGATTTGAGAGAGCGAAGACGAGTGGGGGTGCCCCACTGC	2561
Qy	2521	CTGACCTCTCGGGGCGCAGCGGCACTCATATCGTGGCGCGCAGTTCTCGGTCCGAATG	2580
Dp	2562	CTGACCTCTCGGGGCGCAGCGGCACTCATATCGTGGCGCGCAGTTCTCGGTCCGAATG	2621
Qy	2581	GAGAAAGTGGATTGAGCAATCCAGATGGCCATTGACCTGTGCGAGGAAGCAGCAGCCCC	2640
Dp	2622	GAGAAAGTGGATTGAGCAATCCAGATGGCCATTGACCTGTGCGAGGAAGCAGCAGCCCC	2681
Qy	2641	GCCCCCTAAGTTCTGTGGCAGCAGCCCCCTTGACAAACAAGTCCCTTATGAAAGCAGCGCG	2700

Db	2682	CCCCCTGAGTTCTCTGGCCAGCAGCCCCCTGTACAACAAGTCCCTGATGAAAGCCACGGCG	2741
Qy	2701	GCTGACCAAGAGTCAAGAGATGACTTGAGCGGCTCTGCCACATCGCTGAGGCCAGGCC	2760
Db	2742	GCTGACCAAGAGTCAAGAGATGACTTGAGCGGCTCTGCCACATCGCTGAGGCCAGGCC	2801
Qy	2761	CCGCAACCGGCGCAACAAATGTTGACGTGTGCTGGCAACCCGACACCAAGGTCCTCCATG	2822
Db	2802	CCGCAACCGGCGCAACAAATGTTGACGTGTGCTGGCAACCCGACACCAAGGTCCTCCATG	2861
Qy	2821	GTGGACCTTCAGCATCGCAGTGGAGAAATGATGTGTGAAAACTTCTGAGGAAATTCAAA	2880
Db	2862	GTGGACCTTCAGCATCGCAGTGGAGAAATGATGTGTGAAAACTTCTGAGGAAATTCAAA	2921
Qy	2881	AACAGCAACGGGATGGCAGAAAGCTGTGGTGTTCACAAATCTTCGCTGTCTTCTTAC	2940
Db	2922	AACAGCAACGGGATGGCAGAAAGCTGTGGTGTTCACAAATCTTCGCTGTCTTCTTAC	2981
Qy	2941	AAATTCACACCAAGACAAATCATTCCTTTCGCAAGCTGCTCTGTGGCTACTCGTTCAC	3000
Db	2982	AAATTCACACCAAGACAAATCATTCCTTTCGCAAGCTGCTCTGTGGCTACTCGTTCAC	3041
Qy	3001	ATCCCTCTGAGTCCGAGAAATCATCCAGAAAGACTGACGTTCAGGCTGACCTTCAAGTCC	3060
Db	3042	ATCCCTCTGAGTCCGAGAAATCATCCAGAAAGACTGACGTTCAGGCTGACCTTCAAGTCC	3101
Qy	3061	CACGTCTACTACTTCAGGGCGGAAACGAGTACAGTTCGAAAGGTGATGGAAGTATC	3120
Db	3102	CACGTCTACTACTTCAGGGCGGAAACGAGTACAGTTCGAAAGGTGATGGAAGTATC	3161
Qy	3121	CGCAGTGCACACAGCTCTGCTCCGAGACCCACGTGTGAGCCACAAAGACTCTTGTG	3180
Db	3162	CGCAGTGCACACAGCTCTGCTCCGAGACCCACGTGTGAGCTCAAAAGACTCTTGTG	3221
Qy	3181	TATTGATGGCGGAGCACACTGTTCCGAGGGGTGCTTCCGGAAGAGTTTCCCTT	3240
Db	3222	TATTGATGGCGGAGCACACTGTTCCGAGGGGTGCTTCCGGAAGAGTTTCCCTT	3281
Qy	3241	CTTCTGTAATTATGAGAGCTGTAATAATTAACACTGTCTGAAAAATCAAAACAATGAGCTT	3300
Db	3282	CTTCTGTAATTATGAGAGCTGTAATAATTAACACTGTCTGAAAAATCAAAACAATGAGCTT	3341
Qy	3301	CCGACGAGCTCTCTGTCTCCACAGCGCGGTTTTTAAACCCGACCTCTCAGCGTTGAA	3360
Db	3342	CCGACGAGCTCTCTGTCTCCACAGCGCGGTTTTTAAACCCGACCTCTCAGCGTTGAA	3401
Qy	3361	TGAACAGCGCTCCCAACCTTCAGACTCTGGCAATCCGCTGGGGGCGCTGTCTTACTAGT	3420
Db	3402	TGAACAGCGCTCCCAACCTTCAGACTCTGGCAATCCGCTGGGGGCGCTGTCTTACTAGT	3461
Qy	3421	CCAGTATTAAACAATGTGCTATT 3442	
Db	3462	CCAGTATTAAACAATGTGCTATT 3483	
RESULT 3			
US-10-450-763-564			
Sequence 564, Application US/10450763			
Publication No. US20050196754A1			
GENERAL INFORMATION:			
APPLICANT: Hyseq, Inc			
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES			
FILE REFERENCE: 790CIP3/US			
CURRENT APPLICATION NUMBER: US-10/450,763			
CURRENT FILING DATE: 2003-06-11			
PRIOR APPLICATION NUMBER: PCT/US01/08651			
PRIOR FILING DATE: 2001-03-30			
PRIOR APPLICATION NUMBER: 09/540,217			
PRIOR FILING DATE: 2000-03-31			
PRIOR APPLICATION NUMBER: 09/649,167			
PRIOR FILING DATE: 2000-08-23			
NUMBER OF SEQ ID NOS: 60736			
SOFTWARE: Custom			

	;	SEQ ID NO	564	
	;	LENGTH:	3187	
	;	TYPE:	DNA	
	;	ORGANISM:	Homo sapiens	
	;	FEATURE:		
	;	NAME/KEY:	SIMILAR	
	;	LOCATION:	(49)..(1371)	
	;	OTHER INFORMATION:	92% homologous to Homo sapiens CDEP, accession number AB008430, Smith-Waterman Score=2090.	
	;	OTHER INFORMATION:	AB008430, Smith-Waterman Score=2090.	
	US-10-450-	763-564		
Query Match		92.1%;	Score 3170.2;	DB 10; Length 3187;
Best Local Similarity		99.9%;	Pred. No. 0;	
Matches 3183;	Conservative	0;	Mismatches 3;	Indels 1; Gaps 1;
OY	1	CGCCGCAACGCCCGCGCTGTGGAGATATTCTTAAGCCGCTTTCATCATGGAGAATA	60	
Db	1	CGCCGCAGCCCGCGCGCTGTGGAGATTTCTTAAGCCGCTTTCATCATGGAGAATA	60	
OY	61	GAGCAGAGGCGGACCCCGAGATCAAGACTGGGGGCCCCGGAATAATCGGGATCACTACC	120	
Db	61	GAGCAGAGGCGGACCCCGAGATCAAGACTGGGGGCCCCGGAATAATCGGGATCACTACC	120	
OY	121	TTTGAAACGTGGACAAGAACGGCCCCCAACCTTCAGGAAAACCTGTCATCAAATC	180	
Db	121	TTTGAAACGTGGACAAGAACGGCCCCCAACCTTCAGGAAAACCTGTCATCAAATC	180	
OY	181	CAGATGCTGGATGACACCCAGAGGACATTTGAAGTTCACAAAGAGCTCTGGGAAGGTG	240	
Db	181	CAGATGCTGGATGACACCCAGAGGACATTTGAAGTTCACAAAGAGCTCTGGGAAGGTG	240	
OY	241	CTGCTGGATGACGTTTGTGCAACCACTCAACCTCGTGAAGGTGACTAATTTGGCCTCGAG	300	
Db	241	CTGCTGGATGACGTTTGTGCAACCACTCAACCTCGTGAAGGTGACTAATTTGGCCTCGAG	300	
OY	301	TTTTCTGATCACAAAAAGATCAAGGTGTGCTGGATCTCTTAAACCATTGTGAAACAG	360	
Db	301	TTTTCTGATCACAAAAAGATCAAGGTGTGCTGGATCTCTTAAACCATTGTGAAACAG	360	
OY	361	ATTGAAGAGCCAAAGACAGCTGTGTTGAATTTGNGTGAATTTCTTCCGCTGCACAC	420	
Db	361	ATTGAAGAGCCAAAGACAGCTGTGTTGAATTTGNGTGAATTTCTTCCGCTGCACAC	420	
OY	421	AACAACCTCCAAGAGAACTCAAGAGTACTGTTCGGCTGACAGGTGAACAGAGACTTG	480	
Db	421	AACAACCTCCAAGAGAACTCAAGAGTACTGTTCGGCTGACAGGTGAACAGAGACTTG	480	
OY	481	GCTCAAGGACAGTTGACGTTGAAGTGAACACGACGACGCTCTTGATTTCAACATTTGTG	540	
Db	481	GCTCAAGGACAGTTGACGTTGAAGTGAACACGACGACGCTCTTGATTTCAACATTTGTG	540	
OY	541	CAATCTGAGATTTGGGGATTTTGAAGAACCTTGAACAGAGGCACTTAGCAAAAAATPAA	600	
Db	541	CAATCTGAGATTTGGGGATTTTGAAGAACCTTGAACAGAGGCACTTAGCAAAAAATPAA	600	
OY	601	TACATACCTCAGCAAGACGCACTAGAGGCAAAAATCGTGAATTTCAACATTAACACATT	660	
Db	601	TACATACCTCAGCAAGACGCACTAGAGGCAAAAATCGTGAATTTCAACATTAACACATT	660	
OY	661	GGAACAAACACGACGAAATCAGATTTTCCAGTCTCTTAAGATTTGCCCGCTGAGAGTG	720	
Db	661	GGAACAAACACGACGAAATCAGATTTTCCAGTCTCTTAAGATTTGCCCGCTGAGAGTG	720	
OY	721	TATGGAATTCGGGTGGACCCGGCCCAAGAGACAGGGAAGGACAGAAATCAATCTGCGTT	780	
Db	721	TATGGAATTCGGGTGGACCCGGCCCAAGAGACAGGGAAGGACAGAAATCAATCTGCGTT	780	
OY	781	GCACAACAGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCC	840	
Db	781	GCACAACAGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCC	840	
OY	841	AAGGTGCGGAAGCTGAGCTTCAAGAGAACGCTTTCTCATCAAGCTTCGGCCAGATGCC	900	
Db	841	AAGGTGCGGAAGCTGAGCTTCAAGAGAACGCTTTCTCATCAAGCTTCGGCCAGATGCC	900	

Db 841 AAGGTGCGAAGCTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGCGCAGATGCG 900
Qy 901 AATAGGCGGTACAGAGATACCTTGGAAATCCGTGATGGCAGTCGGGATTTCTGCAATGCC 960
Db 901 AATAGGCGGTACAGAGATACCTTGGAAATCCGTGATGGCAGTCGGGATTTCTGCAATGCC 960
Qy 961 TTCTGGAAATCTGTGTTGAACATCATGCTTCTTTTGAATTGTTTGAAGGCCCAACCA 1020
Db 961 TTCTGGAAATCTGTGTTGAACATCATGCTTCTTTTGAATTGTTTGAAGGCCCAACCA 1020
Qy 1021 AAGCCCAAGCCGCTCTTTTGAAGCGGGGTCACTATTTGGTTCACTGCTGGACTCG 1080
Db 1021 AAGCCCAAGCCGCTCTTTTGAAGCGGGGTCACTATTTGGTTCACTGCTGGACTCG 1080
Qy 1081 AAGCAGGTTCTGCACTATGTTAAAGAGGAGCATTAAGAGGTGAGTTTGAAGAGAA 1140
Db 1081 AAGCAGGTTCTGCACTATGTTAAAGAGGAGCATTAAGAGGTGAGTTTGAAGAGAA 1140
Qy 1141 CACAGCAAGATTCTATTCGAGACCTTGTCTTCAAGCCTTACAGAACTGAATTGGA 1200
Db 1141 CACAGCAAGATTCTATTCGAGACCTTGTCTTCAAGCCTTACAGAACTGAATTGGA 1200
Qy 1201 GTGCTGAGACAGTCTCAGCAGAGCACCGCTTACATTTGAGAGAGAGTGGCGAATTTCC 1259
Db 1201 GTGCTGAGACAGTCTCAGCAGAGCACCGCTTACATTTGAGAGAGAGTGGCGAATTTCC 1259
Qy 1260 AAGGGGCGAAGGTGCGGGGAGAGAAAGAACCGAAGTTTCCGCGGGAGCGCGGGTCC 1319
Db 1260 AAGGGGCGAAGGTGCGGGGAGAGAAAGAACCGAAGTTTCCGCGGGAGCGCGGGTCC 1319
Qy 1320 GCAAGCCGAGCCCTGCGCGGAGAGAAAGCCCGCGGGTAAACAAGCAGCGAGCGAGCCG 1379
Db 1320 GCAAGCCGAGCCCTGCGCGGAGAGAAAGCCCGCGGGTAAACAAGCAGCGAGCGAGCCG 1379
Qy 1380 GCAAGCCGAGCCCTGCGCGGAGAGAAAGCCCGCGGGTAAACAAGCAGCGAGCGAGCCG 1380
Db 1380 GCAAGCCGAGCCCTGCGCGGAGAGAAAGCCCGCGGGTAAACAAGCAGCGAGCGAGCCG 1380
Qy 1381 CTCGAGGCGCCACGAGAGAGAGAGAGAGTGTGTTAAGATAGGACCAAGAGTAAAC 1440
Db 1381 CTCGAGGCGCCACGAGAGAGAGAGAGAGTGTGTTAAGATAGGACCAAGAGTAAAC 1440
Qy 1440 TCAAGCCCGCGAGCCAAAGACAGGCTCCCTGATGAGCTGAGCTTCACTTTCGAGCTGTC 1499
Db 1440 TCAAGCCCGCGAGCCAAAGACAGGCTCCCTGATGAGCTGAGCTTCACTTTCGAGCTGTC 1499
Qy 1500 TGTGAACTGCAAGGGGGAGTGGCCCTGCGCACTGACCTTGTCTCCCAACTGAGCC 1559
Db 1500 TGTGAACTGCAAGGGGGAGTGGCCCTGCGCACTGACCTTGTCTCCCAACTGAGCC 1559
Qy 1560 CGAAGCCAGAGAGGCTCTTCTTGTATCAAGCCGCTGCTGAATGACAGGCTGCCCCG 1619
Db 1560 CGAAGCCAGAGAGGCTCTTCTTGTATCAAGCCGCTGCTGAATGACAGGCTGCCCCG 1619
Qy 1620 GAGGAGCATGAGAGTGAAGGGCGGAGAGAGATTCCCACTGATTAAGCGTACTTCAT 1679
Db 1620 GAGGAGCATGAGAGTGAAGGGCGGAGAGAGATTCCCACTGATTAAGCGTACTTCAT 1679
Qy 1680 AGCTTAAGAAAGTGTCTACACCGAGCGAACTATCTGAAGATCTGAAATTATCACTTC 1739
Db 1680 AGCTTAAGAAAGTGTCTACACCGAGCGAACTATCTGAAGATCTGAAATTATCACTTC 1739
Qy 1740 GTGCTTCAAGAGCATGAGCAAAAGAGAGCGCATGCGGAAAGCATGAAAGTTCAT 1800
Db 1740 GTGCTTCAAGAGCATGAGCAAAAGAGAGCGCATGCGGAAAGCATGAAAGTTCAT 1800
Qy 1800 AATCCCGAATTTTGAACCTTTGCAAAATTTCTAATTTTCTCAAGAAATTGAGA 1859
Db 1801 AATCCCGAATTTTGAACCTTTGCAAAATTTCTAATTTTCTCAAGAAATTGAGA 1859
Qy 1860 AAGACTTGGCGTGTGGAAAGCGCTCAATGCGCAATCAGATTAAGAAATTCG 1919
Db 1861 AAGACTTGGCGTGTGGAAAGCGCTCAATGCGCAATCAGATTAAGAAATTCG 1919
Qy 1920 CGATGTCACTGTAAGAAATTTCAAGGAGTGAAGCATGCGGCTCACTGTGGAAGCA 1979
Db 1921 CGATGTCACTGTAAGAAATTTCAAGGAGTGAAGCATGCGGCTCACTGTGGAAGCA 1979

Qy 1980 CAGGAGGCTTGAAGGCGCTGAGAAATGAATCAAGAGCTCCGCGGCTGGAATCTT 2039
Db 1981 CAGGAGGCTTGAAGGCGCTGAGAAATGAATCAAGAGCTCCGCGGCTGGAATCTT 2040
Qy 2040 CTGCAAGACTTGTGAGCTGCAAGAGGTGTATTCCTACCGCTCAACACTTCTCTCG 2099
Db 2041 CTGCAAGACTTGTGAGCTGCAAGAGGTGTATTCCTACCGCTCAACACTTCTCTCG 2099
Qy 2100 GGCATGCAAGGCTGCAATGCACTAACAAGGTCTGTGAAGCGCTGTGCAACACACCC 2159
Db 2101 GGCATGCAAGGCTGCAATGCACTAACAAGGTCTGTGAAGCGCTGTGCAACACACCC 2159
Qy 2160 GCGGAGCAGCGGCTTCAAGGAGCTGCGAGCCGCTTGTGAGAGATCAAGAGATGAT 2219
Db 2161 GCGGAGCAGCGGCTTCAAGGAGCTGCGAGCCGCTTGTGAGAGATCAAGAGATGAT 2219
Qy 2220 GGCACAGCTTCAAGGAGTGAATCAAGATGAGAGATTTCCAGAACTGCAACTCA 2279
Db 2221 GGCACAGCTTCAAGGAGTGAATCAAGATGAGAGATTTCCAGAACTGCAACTCA 2280
Qy 2280 GAAAGATTGATTTGAGCTTGAATCAATCTGTGTTCCGGGAGGGAGTTCACTCGTGG 2339
Db 2281 GAAAGATTGATTTGAGCTTGAATCAATCTGTGTTCCGGGAGGGAGTTCACTCGTGG 2340
Qy 2340 CAGCTTCAAGAGCTCTGCGGGAGGGGCTCCAGCAGGCACTGTTCTTCTGTTCAACGA 2399
Db 2341 CAGCTTCAAGAGCTCTGCGGGAGGGGCTCCAGCAGGCACTGTTCTTCTGTTCAACGA 2400
Qy 2400 GGTCTGCTATCAAGAGCTGCGGGAGGGGCTCCAGCAGGCACTGTTCTTCTGTTCAACGA 2459
Db 2401 GGTCTGCTATCAAGAGCTGCGGGAGGGGCTCCAGCAGGCACTGTTCTTCTGTTCAACGA 2460
Qy 2460 GGTCTGCTATCAAGAGCTGCGGGAGGGGCTCCAGCAGGCACTGTTCTTCTGTTCAACGA 2519
Db 2461 GGTCTGCTATCAAGAGCTGCGGGAGGGGCTCCAGCAGGCACTGTTCTTCTGTTCAACGA 2520
Qy 2520 CCGTGAAGCTGCGGGAGGGGCTCCAGCAGGCACTGTTCTTCTGTTCAACGA 2579
Db 2521 CCGTGAAGCTGCGGGAGGGGCTCCAGCAGGCACTGTTCTTCTGTTCAACGA 2580
Qy 2580 GGAAGAGTGTGTTGAGAGCATTCAGATGAGCACTTGAAGAGAGAGAGAGAGAGAGAG 2639
Db 2581 GGAAGAGTGTGTTGAGAGCATTCAGATGAGCACTTGAAGAGAGAGAGAGAGAGAGAG 2640
Qy 2640 GCGCCCTGAGTTCGAGGAG 2699
Db 2641 GCGCCCTGAGTTCGAGGAG 2700
Qy 2700 GAGTGAAGTGAAGAGTGAAG 2759
Db 2701 GAGTGAAGTGAAGAGTGAAG 2760
Qy 2760 CCGGAGCGCGGAG 2819
Db 2761 CCGGAGCGCGGAG 2820
Qy 2820 GGTGAGCTTCAAGAGTGAAG 2879
Db 2821 GGTGAGCTTCAAGAGTGAAG 2880
Qy 2880 AAAAG 2939
Db 2881 AAAAG 2940
Qy 2940 CAAATCAACAG 2999
Db 2941 CAAATCAACAG 3000
Qy 3000 CATCCCTCTGAGTCCGAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3059
Db 3001 CATCCCTCTGAGTCCGAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060

QY 3060 CCACGTCTACTTCTGAGGGCGGAAAGCGAGTACACGTTGAAAGGTGATGATGAT 3119
DB 3061 CCACGTCTACTTCTGAGGGCGGAAAGCGAGTACACGTTGAAAGGTGATGATGAT 3120
QY 3120 CCGCAGTGGCCAGCAGCTTGTGCGGACCCACGTTGTAGCCACAAAGTCTTGT 3179
DB 3121 CCGCAGTGGCCAGCAGCTTGTGCGGACCCACGTTGTAGCCACAAAGTCTTGT 3180
QY 3180 GTATTGA 3186
DB 3181 GTATTGA 3187
RESULT 4
US-09-925-297-220
Sequence 220, Application US/09925297
Patent No. US20020081659A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA105
CURRENT APPLICATION NUMBER: US/09/925,297
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05989
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
NUMBER OF SEQ ID NOS: 928
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 220
LENGTH: 3094
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (2)
OTHER INFORMATION: n equals a,c,g, or c
NAME/KEY: misc feature
LOCATION: (4)
OTHER INFORMATION: n equals a,c,g, or c
US-09-925-297-220
Query Match 49.1%; Score 1688.6; DB 3; Length 3094;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1699; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
QY 1740 GTGGTTTCAGAGCAGAGTGAAGAAAGAGAGCGCATGCCGAGACACTGAAAAGTCTCAT 1799
DB 62 GTGGTTTCAGAGCAGAGTGAAGAAAGAGAGCGCATGCCGAGACACTGAAAAGTCTCAT 121
QY 1800 ATTCCCGAATTTGAACTTTGCAAAATTTCTACTTAATTTTCTCAAGAAATTGAGCA 1859
DB 122 ATTCCTCGAATTTGAACTTTGCAAAATTTCTACTTAATTTTCTCAAGAAATTGAGCA 181
QY 1860 AGCACTTGCCTGTGGGAAGCGGCTCAAAATGCCCAATAGAGATTTACCAAAAGATCGG 1919
DB 182 AGCACTTGCCTGTGGGAAGCGGCTCAAAATGCCCAATAGAGATTTACCAAAAGATCGG 241
QY 1920 CGATGTCTATGCTGAAGAACTTCAGGGCATGAAGACCTGGCGGCTCACTGTGAAGCA 1979
DB 242 CGATGTCTATGCTGAAGAACTTCAGGGCATGAAGACCTGGCGGCTCACTGTGAAGCA 301
QY 1980 CAGCGAGGCTTGGAGGCTTGGAGAAATGGAATCAAGAGTCCCGCGGTGAGAACTT 2039
DB 302 CAGCGAGGCTTGGAGGCTTGGAGAAATGGAATCAAGAGTCCCGCGGTGAGAACTT 361
QY 2040 CTSCAGAGACTTTGAGCTGCAAGAGGTGTACTACCGCTCAAGACCTTCTCTGCG 2099
DB 362 CTSCAGAGACTTTGAGCTGCAAGAGGTGTACTACCGCTCAAGACCTTCTCTGCG 421
QY 2100 GCCACTGCACCGGCTCATGCACTAACAAGAGTCTGAGACGGCTGTGCAACCAACCC 2159
DB 422 GCCACTGCACCGGCTCATGCACTAACAAGAGTCTGAGACGGCTGTGCAACCAACCC 481

QY 2160 GCCAGGCCACGGCGAATTCAGAGGAATGCGGAGCGGCTTTGGAGAGATCAGGAGATGAT 2219
DB 482 GCCAGGCCACGGCGAATTCAGAGGAATGCGGAGCGGCTTTGGAGAGATCAGGAGATGAT 541
QY 2220 GGCAACAGCTTCAAGGATGATGATCAAGATGAGAAATTTCCAGAAAGTGCACAGAACTCAA 2279
DB 542 GGCAACAGCTTCAAGGATGATGATCAAGATGAGAAATTTCCAGAAAGTGCACAGAACTCAA 601
QY 2280 GAAAGATTTGATTGGCATTGACATCTTGTGTTCCGGGAAGGAGTTATCCGTCTGGG 2339
DB 602 GAAAGATTTGATTGGCATTGACATCTTGTGTTCCGGGAAGGAGTTATCCGTCTGGG 661
QY 2340 CAGCCTCAGCAAGCTCTCGGGGAAGGGGCTCCAGCAGCGCATGTTCTTCTGTTCAAGA 2399
DB 662 CAGCCTCAGCAAGCTCTCGGGGAAGGGGCTCCAGCAGCGCATGTTCTTCTGTTCAAGA 721
QY 2400 CGTCTCTATACAGAGCGGGGGCTGACGGGCTTCAATCAGTTTAAAGTCCACGGGCA 2459
DB 722 CGTCTCTATACAGAGCGGGGGCTGACGGGCTTCAATCAGTTTAAAGTCCACGGGCA 781
QY 2460 GCTCCCGCTCTATGGCATGACATTTGAGAGACCGAAGCAGAGTGGGGGTGCCCACTG 2519
DB 782 GCTCCCGCTCTATGGCATGACATTTGAGAGACCGAAGCAGAGTGGGGGTGCCCACTG 841
QY 2520 CCTGACCCCTCGGGGCGAGCGGAGTCCATCATGCGGCGGCGAGTCTCGGTCGAGAT 2579
DB 842 CCTGACCCCTCGGGGCGAGCGGAGTCCATCATGCGGCGGCGAGTCTCGGTCGAGAT 901
QY 2580 GGAGAAGTGGTTGAGAGCAGTCAGATGCGCATTTGACCTGGCGGAGAAAGCAGCAGCC 2639
DB 902 GGAGAAGTGGTTGAGAGCAGTCAGATGCGCATTTGACCTGGCGGAGAAAGCAGCAGCC 961
QY 2640 CGGCCCTGAGTTCCTGGCCAGCAGAGCCCTTGAACAAGTCCCTGATGAAGCAACCGC 2699
DB 962 CGGCCCTGAGTTCCTGGCCAGCAGAGCCCTTGAACAAGTCCCTGATGAAGCAACCGC 1021
QY 2700 GGCTGACGAGAGTCAAGAGATGACCTGAGCGGCTTGGGACATGCTGAGACCGCAGGC 2759
DB 1022 GGCTGACGAGAGTCAAGAGATGACCTGAGCGGCTTGGGACATGCTGAGACCGCAGGC 1080
QY 2760 CCGCAGCCGCGCAACACATATGTCACGTGCTGTCGACCCGCAACCCAGCGTCTCAT 2819
DB 1081 CCGCAGCCGCGCAACACATATGTCACGTGCTGTCGACCCGCAACCCAGCGTCTCAT 1140
QY 2820 GGTGAACTTTCAGCATGCGAGTGAAGATCATGTTGCTGAAACCTGCTGAGAAATTCAA 2879
DB 1141 GGTGAACTTTCAGCATGCGAGTGAAGATCATGTTGCTGAAACCTGCTGAGAAATTCAA 1200
QY 2880 AAAACGCAACGGGTGGGAGAACTGTGGGTGTTCAAACTTCTGCTGTTCTTCA 2939
DB 1201 AAAACGCAACGGGTGGGAGAACTGTGGGTGTTCAAACTTCTGCTGTTCTTCA 1260
QY 2940 CAAATACACACAGAGCAATCATCTCCCTGCGAGCTGCTGCTCGCTACTCGTCAAC 2999
DB 1261 CAAATACACACAGAGCAATCATCTCCCTGCGAGCTGCTGCTCGCTACTCGTCAAC 1320
QY 3000 CATCCCTCTGAGTCCAGAACATTCAGAAAGATCAAGTGTCAAGCTGCACTTCAAGTC 3059
DB 1321 CATCCCTCTGAGTCCAGAACATTCAGAAAGATCAAGTGTCAAGCTGCACTTCAAGTC 1380
QY 3060 CCAAGTCTATCACTTCAAGGGCGGAAAGCGAGTACAGTTCGAAAGGTGATGGAAGTAT 3119
DB 1381 CCAAGTCTATCACTTCAAGGGCGGAAAGCGAGTACAGTTCGAAAGGTGATGGAAGTAT 1440
QY 3120 CCGCAGTGGCACAGCTTGTGCTGCGACCCCAAGTGTGAGCCCAAAAGATCTTGT 3179
DB 1441 CCGCAGTGGCACAGCTTGTGCTGCGACCCCAAGTGTGAGCCCAAAAGATCTTGT 1500
QY 3180 GTATTGATGGCCGAGACACTGTTTCCGAGTGGCTGTTCTGGAAGAGCTTCTT 3239
DB 1501 GTATTGATGGCCGAGACACTGTTTCCGAGTGGCTGTTCTGGAAGAGCTTCTT 1560

3240 TCTTGTATTAATGAGCTGTGTAATTAACACCTGTCTGTAATCAAAAAATCATGCT 3299
1561 TCTTGTATTAATGAGCTGTGTAATTAACACCTGTCTGTAATCAAAAAATCATGCT 1620
3300 TCCGACAGCTCTCTGTCTTCCACAGCCGCGTTTAAACCCGACCTCTGAGCTTTGA 3359
1621 TCCGACAGCTCTCTGTCTTCCACAGCCGCGTTTAAACCCGACCTCTGAGCTTTGA 1680
3360 ATGAACAGGCTCCGACCTCCAGTCTGAGATCCGCTGGGGGCGCTCTTTAGCTAGT 3419
1681 ATGAACAGGCTCCGACCTCCAGTCTGAGATCCGCTGGGGGCGCTCTTTAGCTAGT 1740
3420 GCCAGTATTTAAACATTTGCTATT 3442
1741 GCCAGTATTTAAACATTTGCTATT 1763

RESULT 5
US-10-172-118-1536
Sequence 1536, Application US/10172118
Publication No. US20030224374A1
GENERAL INFORMATION:
APPLICANT: Dai, Hongyue
APPLICANT: He, Yudong
APPLICANT: Lineley, Peter
APPLICANT: Mao, Mao
APPLICANT: Roberts, Chris
APPLICANT: Van 't Veer, Laura
APPLICANT: Van de Vijver, Marc
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-175-999
CURRENT APPLICATION NUMBER: US/10/172,118
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 60/380,770
NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 1536
LENGTH: 3997
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: NM_014808
DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1536

Query Match 30.1%; Score 1037.4; DB 7; Length 3997;
Best Local Similarity 60.2%; Pred. No. 1.1e-305;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

69 GCGACCCGAGATCAAGACTGGGGGCCCCGGAATTCGGGGATCAGTACCTTGGAAAG 128
150 GCGAGCTGAGGGAGTGGCTGGGTGCCAGAGCCCTGTGGAGTTAGACCTTGAGCC 209
129 TGGACGAAGCCGCCCCCAACCTTCAGGAAAATCTGTCTCATCAAAATCCAGATGCT 188
210 TGGGCGAGACTCTCTTCCAGAAATGCAAGAAAGCACTGACCTCAGATGAAGCTGCT 269
189 GAGTGCACCCAGAGAGCATTTGAAGTTCACAAAAGAGCTCCCTGGGAAGTGTCTGGA 248
270 GGAACAACCATGGAATATTTGACATTTAGCTTAATGCGAGTGGCAGGTATTTATGAC 329
249 TGCAGTTTCAACCACTCAACCTCTGGAAGGTGACTATTTTGGCTGAGTTTCTGA 308
330 ACAAGTGTGAGAGCTTTAACTGTGTAATGTGACTACTTCGGGATGGAATTTCAAA 389
309 TCACAAAAGATCACGCTGTGCTGATCTCTTAAACCCATTTGTAACAGATTAGAG 368
390 TACTGATCTACTGATTTGGCTTGAACCTATGAACCAATCATTAAGCAATAGAG 449
369 GCGAAGGACATTTGTTTGAAGTTTGTGTGAATTTCTTCCGCTGACACACAACT 428
450 GCCAAAGATGTGTGCTTGTGCTTGTAGCTGTAAATTTTTCACCTGATCTGTGTAGCT 509

429 CCAAGAAAGTCAAGAGTACTGTTCGCGCTGAGGTGAAGCAGACTTGGCTCAAG 488
510 ACAGAAAGATATACAGATATCTTGTTCCTTGCACTTAAGAGAGCTGTGGAGA 569
489 CAGTTGACGTGTAAATGACACACGCGCAGCTCTCTTGAATTTCAACATTTGGCAATCTGA 548
570 GCGTTGACCTGTGTGACACACAGCGGCTCTCTCAGCTCCATCTCTGCACTGCGA 629
549 GATTGGGATTTTGTATGAAGCTTGGACAGAGAGCACTTACGAAAATTAATATATCC 608
630 ATATGAAGATTTACGATGAAGAGCTGAGACCGAGAGCACTCAAGTGAACGAGATTTGGC 689
609 TCAGCAAGACGACATAGAGCAAAATGTGGAATTTTCCATTAACCAATTTGACAAAC 668
630 TGGCAGAGCAGCTGCTTGAAGATATCTGAATTTCAATGAAAGCAGTGGGCGAGC 749
669 ACCAGCAGATCAGATTTTCCAGCTCTTGAAGATTTGCGGCTGAGATGTATGAAT 728
750 ACTGCTGAGTGTGATTTTCCAGTGTCTGAAATTTGCTGAAAGTTGGAATGTACGCGAT 809
729 CCGGTTGACCCGCGCAAGGACAGGGAAGGACAGAAATGCAATCGGCGTTGCCAAC 788
810 CAGATTTGACATGTGCTTCTGACAGGGAAGAAACCAAGATTCACCTGAGATTTCCACAT 869
789 GGGATTTCTAGTGTTCAGGCTTTCATAGATCAATGCTTCAACCTGAGGCGCAAGTGTG 848
870 GGGTGTACTCGTGTTCAGGCGGACACCAAAATCAACCTTCAACCTGATCCAGTCCG 929
849 GAGCTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGCGCAATGCCAATATGTC 908
930 TAAATTAAGCTTCAAGAGAAAGATTTCTTATCAATCACTTCAACAGGTTCAAGACC 989
909 GTACAGAGATCCTTGAATTCCTGATGAGGAGTGGGATTTCTGCAAGTCTCTTGTGAA 968
930 TTACAGAGACATATGAAATTTGTGTGATGATGATGATGATGATGATGATGATGATG 1049
969 AATCTGTGTGAACATCATGCTCTCTTGTGACCTTTTGAAGAGGCCAAACCAAGCCCA 1028
1050 GATTTGTGTGAGATTCACACCTTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1109
1029 GCGGCTCTCTTGTAGCGGGGCTCATCTTTCGTTGCTGATGTGAGATCTCAGAGCAGT 1088
1110 AGCCGCTCTTCTTACGCGGGGCTCTCTTCAAGATACAGTGAAGAACTCAGAAACACT 1169
1089 TCTGACTATGTTAAAGAGAGGACATAGAGAGGTGAGTTTGAAGAGAGAGCAGCA 1148
1170 AGTAAATTTATTTCAAGAGCAGTGAATGAGATGAGATTTCAATGTAAGAGAGCAGCA 1229
1149 GATTCATTTCTATCCGAGCTTGTCTTCAAGCTTACAGACTTACAGAACTGAAATGCTGGA 1208
1230 GACCCACAGCTC-----GTTGAGCTCTGACTGACAGCTTCCAA 1271
1209 GCAATCTCAGAGAGCAGCAGCTTACATTTTGAAGAGAGTGTCCGAAATCTCCAGGGGCA 1268
1272 ACAGAGATCTCATTTCCCGAGGAGATTGAGACCTCTTCCATCTTCCAGCAATGCT 1331
1269 GAGCTGCGGCGAGAGAAAGAACCGAAGTTTCCGCGGGAGCGGGGATGCAACCCGAG 1328
1332 CTTTACTGTGCTCTTCCCTCAGCTGTGTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1391
1339 CCCTGCGCGAGAGAGAGCCCGCGGATTAACAAGAGGCGGACGAGCGCTCTGCGGCT 1388
1392 CAGCAGCTCCCTCAGAGATCCCGAGGTTTCCAGCTCAAGAGATGCAAGCTGAGAGAGG 1451
1389 CACGAGAGAAAGAGAGAGG---TGTATTAAGATAGAGCCAGCAGAGATTAACCTCAGCC 1445
1432 CAGTGAACAGATGTGTGAGAGCGCCGACACACATGCGCCCAAGCCCTCGGGCCCCCGC 1511
1446 CCGGAGCGAAGCAGAGCTCCCTGACTGAGAGTCTCAGCTTTCAGAGCTGTCTGTGA 1505
1512 ACTCAGGCTGTCCAGGCTTTTCCAGAGAGATCTTCAAGCTTCTCTCTCTCAGGCGAA 1571

1506 CTGCAAGGGGGAGTGGCCCTTCGCAACGTAACCTTGTCTCCCAACTGAGCCCGACAC 1565
1572 GAGCCCCCTGAGTCTAGAGCCCTGATTTAGGTGCTT-----TGGGCCAGCTGA 1622
1566 CAAGCAGGCTCTCCCTTGAATGAGCCGCTGTGAATGACAGAGCTGCCCGGACGGA 1625
1623 ACAGGGCTATCCCACTCTGTAGCCCTGTCTCAGTGAATGCTGGGAGACCGGAGTGA 1682
1626 CAGTAGAGTGAAGGGCGGAGAGAGATTTCCCACTGAATGAACGCTACTTCTAGCTAA 1685
1683 CTGCGAG---GAGCCAGACACAAAGGCGCTGTGAGAGAGAGAGCCCTTACTTCAATGCA 1739
1686 GGAAGTGTACACCGAGGGAACATATCTGAAGATCTCGAAGTTATCACTTCTGTGTT 1745
1740 AAGAGTTCTGCTACAGAGAGAACTACTCGAAGATTTGAAGTTATTCAGTGTGTT 1799
1746 TCAGACACAGTGAAG 1805
1800 CCGAGAGAGTGTGAAG 1859
1806 GAATTTGAACCTTGTGACAAATTTTCACTAATTTTCTCAAGAAATGAGCAAGACT 1865
1860 CAACATCGATCCCATGTATGAGATTTCAAGAGGCTTCTGCGAGAGTGAAGAGAGGCT 1919
1866 TGCCCTGTGAG 1922
1920 GGCACCTGTGAG 1979
1923 TGTATGTCTGAAG 1982
1980 CATCTGTCTAG 2039
1983 CGAGGCTGTGAG 2042
2040 CAGAGGCTCTGAAG 2099
2043 CAGAGAGCTTGAAG 2102
2100 CAGAGAGTTGAAG 2159
2103 ACTGCAACGAGGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2162
2160 CATTCAGAGGCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2219
2163 GAGCCAGGCTGTCTGAG 2222
2220 CGGGAG 2279
2223 ACAGCTCCAGGAG 2282
2280 CACATTAACAG 2339
2283 AGATTTGATTTGAG 2342
2340 GGAAGCTGTGTGAG 2399
2343 CCTCAG 2402
2400 CTTTCAAG 2459
2403 CTTGATTAACAG 2462
2460 GTTGTCTGTACAG 2519
2463 CCGCTCTATGAG 2522
2520 TCCCTCTCAAG 2579
2523 GAGCCCTCGGAG 2582
2580 CACCATCTACGAG 2639
2583 GAAGTGGTTGAG 2642

2640 GAAAGTGAATGCTGAG 2699
2643 CCTGAGTCTCTGAG 2696
2700 CCTGAG 2759
2697 CCGGAGCTGACAG 2756
2760 TCTG---GAGCAG 2816
2757 GAGCCCGCAG 2816
2817 TGGCCAG 2876
2817 CATGTGAGCTTACAG 2876
2877 CAGGAG 2936
2877 CAAAAAG 2936
2937 CAAAAAG 2996
2937 CTACAAATGACAG 2996
2997 CTACAAATGACAG 3056
2997 CACATCCCTCTGAG 3056
3057 GAGCATCCCGAG 3116
3057 GTCCAGCTCTACTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3116
3117 ATCCAGCTCTACTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3176
3117 GATCCGAG 3176

RESULT 6
US-10-342-887-1536
; Sequence 1536, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberte, Christopher J.
; APPLICANT: Van de Veer, Laura Johanna
; APPLICANT: Van de Veer, Marc J.
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342, 887
; PRIORITY FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298, 918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380, 710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172, 118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1536
; LENGTH: 3997
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1536

Query Match 30.1%; Score 1037.4; DB 8; Length 3997;
Best Local Similarity 60.2%; Pred. No. 1.1e-305;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

QY GCGAAGCCAGAGATCAAGACTGGGGGCCCGGAAAATTCGGGGATCAGTACCTTGGAAAG 128
 DB GCGAGACTGAGAGATGCGCTTGGGTGCCAGAGACCCCTGTGGAGATTAGACCTTGGAGCC 209
 QY TGGACAGAAAGCCGCCCAACACTTCAGAGAAAATCGTGTCCATCAAAATCCAGATGCT 188
 DB TGGGAGACTCTTCTTGTCCAGAAATGCAAGAGAAAGCACTGACCTCAGATTAAGCTGTCT 269
 QY GATATGACCCAGAGAGAGATTGAAATTCACAAAAGATCCCTGGGAAAGTGTGTGTGA 248
 DB GAGCAACACATGGAATATTGACATGAGCTTAATGCGAGTGCAGAGATTACTGAC 329
 QY TGCAGTTTGCACACACTCAACCTCGTGAAGGTGACTATTTTGGCCCTGAGTTTCTGA 308
 DB ACAGGTGTGAGAGCGTTTAAACCTGTGAGAAATGTGACTACTTCGGGATGGAAGTTTCAAA 389
 QY TCACAAAAGATCAAGGTGTGTGATCTCTTAAACCCATTTGTGAACAGATTGANG 368
 DB TACTCAGTCTTACTGATTTGGCTTGAACCTATGAAACCATCATTTAGGCAAAATACGAAG 449
 QY GCGAAGACGTTTGTGTTAAGTTTGTGGAATCTTTCGGCCCTGACACACACAACT 428
 DB GCGAAAAGATGTGTGCTTCTGCTAGCTGTAAATTTTTCACCTGATCTGTGTAGCT 509
 QY CCAAGAGAACTCAACAAGTACCTGTTCCGCTGAGTGAAGCAGAGACTTGGCTCAAG 488
 DB ACAAGAGAAATATACAGATATCTTGTTCCTTGCATTAAGAGAGACTGTGAGANA 569
 QY CAGGTGACGTGTATGACACACAGCGACGCTCTTGTATTTCAACATTTGTGCAATCTGA 548
 DB GGGTTGACCTGTGTGACACACAGCGGCCCTTCTCAGTCCCATCTCTCGAGTCGA 629
 QY GATTGGGGATTTTGTGAAGCCTTGTGACAGAGAGCACTTACGAAAATTAATATACATACC 608
 DB AATAGAGATTAAGATGAACCGTGAACGAGAGCACTCAAGTGAAGAGATTTGCTC 689
 QY TCAGCAAGACGCACTGAGAGCAAAATCGTGAATTTCAACATACACATTTGACAAAC 668
 DB TGGCCAGAGCACTGCTTGAAGAGATTAATGAAATTTCCATCAGAAAGCACTGGGCCAGAC 749
 QY ACCAGCAAGATCAAGATTTCCAGCTCTTGAAGATTTGCCGTGGCTAGAGATGTAGAAAT 728
 DB ACTGTGTGTCGGAATTTCCAGGTGTCTCAAAATGTCTCAAAAGTTGGAATGTACCGCAT 809
 QY CGGTTGCAACCGGCGCAAGAGACAGGGAGGCAAGAGATCAATCTGGCCGTTGCCAACAC 788
 DB CAGATTTCAATGCTTCTGACAGGAGAGAAACCAAGATTTCACTGGCAGTTTCCCAT 869
 QY GGGAAATTCAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCCAAGGTGCG 848
 DB GGGGTGACTGTGTTCAGGGGACACACAAATTCACATTTCACTGGTCCAAAGTCCG 929
 QY GAAAGTGAAGCTTCAGAGAGAGAGCGCTTTCATCAAGCTTCGGCCAGATGCCAATGTGCG 908
 DB TTAACCTTAAGCTTCAGAGAGAGAGAAAGATTTCTTATCAAACTTCACTCCAGAGTTCACTGACC 989
 QY GTACACAGATTAACCTTGGAAATTCCTGATGCGCAGTGGGGAATTTCTGCAAGTCCCTTGGAA 968
 DB TTACACAGAGCACTTAAGAAATTTTGTGGGTAGTAAAGATTAAGATTAAGAACTTTGGGAA 1049
 QY AATCTGTGTGAACATCATGCTTCTTTTGAAGCTTTTGAAGGCCCAACCAAGCCCA 1028
 DB GATTTGTGTGAGATACACACCTTTTGAAGCTTTTGAAGCAACCTTAACCAAAAGCAAA 1109
 QY GCGCCGTCTCTTTAGCGGGGGTCACTTTGCGTTCAAGTGTGCGACTGAGAGAGAGCT 1088
 DB AGCCGCTTCTTCAAGCGGGGCTCTCTCTTCAAGATCAAGTGAAGAACTTCAAGAACACT 1169
 QY TCTCGACTATGTATTAAGAGAGAGACATTAAGAGGCAAGTTTGAAGAGAGAGACAGAA 1148
 DB AGTAGTTATTTCAAGACAGTGAATGAAGAGAAATTCATATGAAGAGAGACAGCA 1229

QY GATTCATTCTATCCGAGACCTTGTCTTCAAGCTTACAGAACTGAATTCGAAAGTCTGGA 1208
 DB GAGCCACAGTCTC-----GTTGAGGTCTGACTGACAGACTTACCAAA 1271
 QY GCAATCTCAGACAGACACACAGCTTTACATTTGGAGAAAGTGTCCGAATCTCCAGGGGCCA 1268
 DB ACAGAGCATCTCATTTCCCGAGAGAGATTGAGGACTTCCTCCTCCCATCTTTCAGCAATGTC 1331
 QY GAGCTGCCGCGAGAGAAAGAAACCGAAGTTTTCGCGCGGAGAGCGGGGTGCAACCCGAG 1328
 DB CTTTACTGTGCTCTCTCCCTCCACTGTGTCTCCCTCTGTGGCTGCGCAGAGTTTAAAGACAG 1391
 QY CCTGCGCGAGAGAGAGAACCCCGCGGTGAACAGAGAGCGGAGCGGCGCTCGCGCC 1388
 DB CAGCAGCTCCCTCACAAGATTCAGAGTTTCTTAAGTCAAGAGTCCAGCTGACAGAGAGCG 1451
 QY CACGAGAGAAAGAGAGAGAG---TGTTTAAAGATTAGACCCACAGCAGATTAACCTCAGCC 1445
 DB CAGTGAACAGTGTGTGAGAGGCCCGGACACACATGAGCCCAAGCCCTCGGGCCCGCC 1511
 QY CCGGAGCGCAAGCAAGGCTCCCTGACTGGCAGTCCCTCAGCTTTCGAGCTGTCTGTGA 1505
 DB ACTCAGCTGTGTCCAGGCTTTTCCAGAGAGATCTCAGCTTTCCTCCACAGCGGAA 1571
 QY CTGCGAGGGGGAGTGTGACCCCTGCGCAAGCTGACTTGTCTCCAACTGAGCCCGGACAC 1565
 DB GAGCCCTGAGTGTGAGCCCTGCACTTCAGGTGCTT-----TGGGCCAGCTGA 1622
 QY CAGCAGAGCTCTCCTCTTGAATCAGCCCGCTGTGAATGAACAGGCTGTGCCCGGACGGA 1625
 DB ACAGGGCTCATCCCACTCTGAGGCCCTGTCTCAGTATGTGTGGAGCGGAGATGGA 1682
 QY CGATGAGATGAGGGCCGAGAGAGATTCCTCACTGAATTAAGGATCTTCAATGACTTA 1685
 DB CTGCGAG---GAGCCACAGACAGAGCGGTCTGTGACAGAGGCTTCTCAATGTGCA 1739
 QY GGAAGTGTCTACACCGAGCGAATATCTGAAGGATCTGAAAGTTATCATTTGCTGTGTT 1745
 DB AGAGATTTCTCGCTACAGAAACAACTACTCAAGGATTTGAAGTTTACCGTGTGTT 1799
 QY TCAGAGCAAGTGAAG 1805
 DB CCGAGCGCAAGTGTGAAG 1859
 QY GAAATTTGAACCTTTGCACAAAATTCATTAATTTTCTCAAGAGAAATGGAACAGACT 1865
 DB CAACTCATATCCATCTATATGAGTTCCAGAGGCTTCTGTGGAGAGTGAAGAGGCT 1919
 QY TGGCTGTGGAGAGGCGGCTCAAAATGCCCCAATCAGAG---ATTACAAAGATTCGCGGA 1922
 DB GGCATCTTGGGAAAGGGCCCTCAAAAGCCCAACAAAGGAGATCATCAACGAATCGGGGA 1979
 QY TGTATGTCTGAAGAACATTTCAAGGAGATGAAGCACTGTGGCGGCTCACTGTGAAGACAG 1982
 DB CATCTGTCTCAGAGAACATGCGCCAGTTAAAGAGTTTACAGCTATCTTCAAAAGACATGA 2039
 QY CGAGGCTTTGAGAGGCGCTGAGAAATGAATCAAGAGCTCCCGCGGCTGAGAACTTCTG 2042
 DB CGAGGTCTTAACAGAACTGGAAGAGGCTTACCAAAGCTGTGAAGAGTTGAGAGGAGTGA 2099
 QY ACTGACCGGCTCATGCACTAACAGAGTCTTGAAGGGCTGTGCAACCAACCGGC 2162
 DB CATTCAGGGGCTGTGCACTACCGCTGTGTGCGCGCTTATGCGGACATTAACAGCCC 2219
 QY GAGCAGCGCCGACTTCAGGAGCTGCGAGCGCTTTTGGCAGAGATCAGAGATGTGTGC 2222
 DB CGGGACACATGACTAGCGTGAAGTGAAGCGCTTGAAGGCACTACAGAGGTACAC 2279
 QY AAGCTCCAGGTAAGATGATCAAGATGAGAAATTTCCAGAAAGCTGACGAACCTCAAGAA 2282

Db 2280 CACACTACAGACATTTCTATCCGGCTGGAGAACCTGCAAGACTAAGAGCTGACGGC 2339
Qy 2283 AGATTGATTGGCATTTGACATCTTTGGTTCCGGGAAGGAGTTCACTCGTGGGAG 2342
Db 2340 GGACTGTGGGATAGAGAACTCATTTGCTCTGGGAGGAGTTCACTCGTGAAGGCTG 2339
Qy 2343 CCTCAGCAAGCTCTCGGGGAGGAGGCTCCAGAGCGCATTTCTCTCTGTTCAACAGCT 2402
Db 2400 CTTTCAAGACTCAACAGAGGAGCTGACAGAGAGATTTTCTGTTCTCAAGTAT 2459
Qy 2403 CTTGCTATACAGAGCGGGGCTGACGGCTCCATCACTTTAAATGTCACAGGAGCT 2462
Db 2460 GTTGCTGTACAGAGGAGGAGTTGAGGAGCAAGCACTTCGGATCCGGGGCTCCT 2519
Qy 2463 CCGGCTCTATGAGTACCAATTGAGAGAGAGCAAGAGTGGGGGCTCCCACTGCT 2522
Db 2520 TCCCTCTCAAGAGATCTGTGTGAAGAAAGTATACAGAGTGTCTGTTCAACATGTTT 2579
Qy 2523 GACCTTCCGGGGCCAGCGGAGTCAATCATGTGTGCGCCAGTTCTCGGTCCGAGATGA 2582
Db 2580 CACACTTACCGCGCTCAAGAAACATCTGTGTGAGAGCCAGCACTCGCTGAGAAAGA 2639
Qy 2583 GAAGTGGTTGAGACATCCAGATGGCCATTGACTGGCGAGAGAGAGAGAGAGAGCCCGC 2642
Db 2640 GAAGTGGATGTGAGCTGAACTCCGGATCCAGAGCCAGAGAGTGGGGGTGACAGCGC 2659
Qy 2643 CCTGATCTCTGCGCAGCA-----GCCCTTCAACAGATCCCTGATGAGAGCCAC 2696
Db 2700 CCTGTGACGCGCCAGCGGAGCTGTGTGCACTGTCTCCCGCAGATCCCGCAACAGAGTATC 2759
Qy 2697 CGCGGCTGACAGAGATGAGAGATGACCTGAGCGCTCGCGCACTCGTGGAGGCGCA 2756
Db 2760 TCTG---GAGCAGAGTCAAGAAATGATGCTCGGGGTGTCGCACTCCCTGAGAGGCA 2816
Qy 2757 GCGCCCGCAGCGCGGCAACAAATGATGCACTGTGTGCGACCGCAACAGAGCTCTC 2816
Db 2817 TGGCCAGACCGGGCCAAACCAATGACGTGTGTGTGATCCGGAACAGAGCGTGT 2876
Qy 2817 CATGTGATCTTACAGATGCGAGTGAAGAAATCAGTTGTCTGGAACCTGTGAGAAAT 2876
Db 2877 CAGGGGAGACCAAGTGAAGCTGTGCAAGAACAGCTTTCAGATATCTGTGAAGAAAT 2936
Qy 2877 CAAAAACAGCAACGGGTGAGAAAGCTGTGGGTGTTTCAAACTTCTCGCTGTTCT 2936
Db 2937 CAAAAACATCATGTGTGAGAAAGCTGTGGGTGTTTCAAACTTCTGTTGTTCT 2996
Qy 2937 CTACAAATCAGACAGAGCAATCATCCCTTGCAGCTGCTCTGCTGAGTACCTGCT 2996
Db 2997 CTACAAATCAGATCAGAGTGAATCAACCATGTGCGCACTCCGCTGTGGGCTACAGCGT 3056
Qy 2997 CACATCCCTCTGAGTCCGAGAAATCCAGAAAGTACGTGTCAAGTCACTTCAA 3056
Db 3057 GAGCATCCCGAGGAGGCGAGTATACAAAGATATGTTTCAAGTCCAGTTCAA 3116
Qy 3057 GTCCCAAGTCTACTTCTAGAGGCGAGAGCGGTCAAGTTCGAAAGGTGATGAAGT 3116
Db 3117 ATCCCAAGTCTACTTCTCGGGCTGAGAGCAAGTACATTTGAAAGGTGATGAAGT 3176
Qy 3117 GATCCGAGTGCACCAAGCTTGTGCTGCGAGCC 3151
Db 3177 GATCCAGGGGGCCAGAGCTCAGCCGGAGGGGCC 3211

RESULT 7
US-10-802-432-16
; Sequence 16, Application US/10802432
; Publication No. US20040185489A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann-La Roche Inc.
; TITLE OF INVENTION: Transcriptional Activity Assay
; FILE REFERENCE: 21574
; CURRENT APPLICATION NUMBER: US/10/802,432

; CURRENT FILING DATE: 2004-03-17
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 3997
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: KIA0793
; LOCATION: (1)-(3997)
; OTHER INFORMATION: GenBank accession No. AB018336
US-10-802-432-16

Query Match 30.1%; Score 1037.4; DB 9; Length 3997;
Best Local Similarity 60.2%; Pred. No. 1,1e-305;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

Qy 69 GCCGACCCAGATACAGACTGGGGGCCCGGAGAAATTCGGGATCAGTACCTTGGAAAG 128
Db 150 GCAGACTGACAGGGATGCGCTTGGGTGCCAGACCCCTGTGGAGTTAGCACCTTGAGCC 209
Qy 129 TGGACAGAAAGCCGCCCAACACTTCAAGAAATCTGTGTCATTAATAATCAATGCT 188
Db 210 TGGGCACTCTTGTCCAGATGCAAGAGAGCACTGCACTCAGATTAAGTGTCT 269
Qy 189 GATGACACCCAGAGGAGCATTTGAAGTTCCACAAAGAGCTCCGGGAGGTGTGTGGA 248
Db 270 GAGCAACACCATGAAATATTTGACATGACCTTAATGCAATGGCAGATTTACTGAC 329
Qy 249 TCGAGTTTCAACCACTCACTCACTCGTGAAGGTGATATTTTGGCTCGAGTTTCTGA 308
Db 330 ACAAGTGTGAAGCGTTTAACTGTGTGAATGTGACTTCTGGAGTGAAGTTTCAAA 389
Qy 309 TCACAAAAAGATCAGCGTGTGTGATCTCTAAACCATTGTGAACAGATTGAAG 368
Db 390 TACTCAGTCTTCTGATTTGCTTGAACCTATGAAACCATATATGAGCAATACGAAG 449
Qy 369 GCCAAGCAAGTTTGTGTTAAATTTGTGTGAAATCTTTCGGCTGACACACCAACT 428
Db 450 GCCAAGAAATGTGTGCTTGGCTTGAATTAATTTTTCACCTGATCTGTGTGCT 509
Qy 429 CCAAGAAAGATCAGAGTACCTGTCGCTGAGGTGAAGAGAGAGCTTGTCAAG 488
Db 510 ACAAGAAATATACAGATATCTTGTGTGCTTGAATTAAGAGAGAGCTGTGAGAA 569
Qy 489 CAGTTGACGTGTATGACACAGAGCAGCTCTTGAATTTCAACATGTGCAATCTGA 548
Db 570 GCGTTGACCTGTGTGACACACAGCGGCCCTTCTCAGTCCCATCTCTCAGTCGGA 629
Qy 549 GATGGGAGTTTGTGAAGAGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 608
Db 630 AATAGAGATTAAGATGAAGAGCTGAGCCGAGAGAGAGAGAGAGAGAGAGAGAG 689
Qy 609 TCAGCAAG 668
Db 690 TGGCAG 749
Qy 669 ACCAGAGATATGATTTTCAAGTCTTGAAGATTTGCGCTGAGAGATGATGAAGT 728
Db 750 ACTGTGTGTGAGATTTTCAAGTGTCTGAAATTTGTCGAAAGTTGAGATGAGAGAG 809
Qy 729 CCGGTTGACCGGGCCAG 788
Db 810 CAGATTTCAATGAGCTTCTGACAGAGAGAGAGAGAGAGATTAAGTGTGAGATTTCCAGAT 869
Qy 789 GGAATTTCAATGATTTTCAAGAGGTTTCACTAAGATGATGCTTCAACTGGGCCAGAGTGG 848
Db 870 GGGTGTACTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 929
Qy 849 GAAGTGAAGTTTCAAG 908
Db 930 TAAACTAAGCTTCAAG 989

QY	909	GTACCAAGATACCTTGGAAATTCCTATGATGCCAGTCGGGATTTCTGCAAGTCCTTCTGGAA	968
Db	990	TTACCAAGACACATTAGAAATTTTGTGGGTAGTAGAATGATAGAACTTCTGGAA	1049
QY	969	AATCTGTGTGAACATCATGACCTTCTTTAGACTTTTGAAGGCCCAACCAAGGCCAA	1028
Db	1050	GATTTGTGTGGAGTATCAACCTTTTATAGACTTTTGAACCAACTTAAAGCCAAAGCAA	1109
QY	1039	GCCCGTCCTCTTTTAAACCGGGGGTCACTTTCGTTTCAGTGATCGGACTCAAGAACAGT	1088
Db	1110	AGCGGTCTTCTTCAACCGGGGGTCTCTCTTGATGATAGTGGAAAGAACTCAAGAAACAAT	1169
QY	1089	TCTGCATCTATGTTAAAGAAAGAGACATTAAGAGTGCAGTTTGAAGAAACACAGCA	1148
Db	1170	AGTAGATTATTTCAAAAGACAGTGGAAATGAAGAAATTCATATGAAAGAAAGCACAGCA	1229
QY	1149	GATTCATTCTATCCGGAGGCTTGCTCAACGCTACAGAACTGAAATGGGAAGTCTGGA	1208
Db	1230	GAACCAACGTCC-----GTTGAGCTCTGACTGACAGCTAACAA	1271
QY	1209	GCAGTCTCAGCAGACACCAGCTTACATTTGGAAAGGTGCGGATCTCCAGGGGGCCA	1268
Db	1272	ACAGAGCAATCTCATTTCCCGAGGGATTAGAGACTCTGTGCTCTCCCATTTCAAGGAATGC	1331
QY	1269	GAGCTGCGGGCGAAGAAAGAAACCGAAGTTTCCGCGGGGAAGCCGGGGTGCACCCGAG	1328
Db	1332	CTTTTACTGCTCTCTCCCTCCACTCTGGTCCCTCTGGGCTGCGCAGAGTTTAAAGACAG	1391
QY	1329	CCCTGCGCGAAGAAAGCCCGGGGTAAACAAGAGCGGACGGAAGCCGCTGGGGCC	1388
Db	1392	CAGCAGCTTCTCAACAGATCCCCAGGTTTCTACGTCAAGATCTCAGCTGACAGAGGCG	1451
QY	1389	CACGAGGAAGAGAGAGAGS--TGGTTAAGATTAGAACCCGACAGATAAACTCAGCC	1445
Db	1452	CAGTGGACAGAGGTGCTGAGAGGCCCGGACACACATCGGCCCAAGCCCTCGGGCCCCCGC	1511
QY	1446	CCCCGAGCCAGACAGAGCTCCCTGACTGCGAGTCTCACTTTTCGAGCTGTCTGAA	1505
Db	1512	ACTCAGGCTGGTCCAGGCTTTTCCAGAAAGATCTCAGCTTCTCCCTCAGCCGGAA	1571
QY	1506	CTCGAAGGGGGAGTGGGCCCTCGCCAAAGTACCTTGTCTCCAACTGAAAGCCCGAAC	1565
Db	1572	GAGCCCCGTGATCTGAGCCCTGCAATTCAGGTGCTT-----TGGGCCCAAGCTGA	1622
QY	1566	CAAAGAGGCTTCTCCCTTGATCAGCCGCTGTGATGATCAGAGGCTTCCCGCGAAGCGA	1625
Db	1632	ACAGGGCTCATCCCACTCTGAGGCCCTGTCTCACTGATATGCTGGCGAAGCCGGGAATGA	1682
QY	1626	CGATGAGATGAGGGCCGGAAGAGATTCCTCAATGATGATGATTAAGCGTATCATAGCTAA	1685
Db	1683	CTGCAGAG--GAGCCCAAGACAAAGCGGTGCTGCAAGAGGGCTTCTCATATGCTAA	1739
QY	1686	GGAATGTCTACACCGAGCCGAATATCTGAAGATCTCGAAGTATCACTTGTGTGT	1745
Db	1740	AGAGATTTCTCGTTACAGAACGAACATCACTCAAGATTTTAAAGTATTTTACCGTGTGTT	1799
QY	1746	TCAGAGCACAGTGAACAAGAGAGAGCGCCATGCGGGAACCTGAAGAAATGTCATATTTCC	1805
Db	1800	CCGACGCGCAGTGTGAAGAGAGACGCACTGCTGCACTTGATGACGCTGCTTCTTC	1859
QY	1806	GAATTTGAACCTTTGCACTAATTTTCTAATCTAATTTTCTCAAGAAATTTGAACAGACT	1865
Db	1860	CAACATCGATCCCATCTATGATTCACAGAGGCTTCTGGCGAGGTGAGAGAGAGGCT	1919
QY	1866	TGCCCTGTGGGAAGGCCGCTCAATGCGCCAATCAGAG--ATTACCAAGATTCGGCGA	1922
Db	1920	GGCATCTTGGGAAGGGCCCTCCMAAGCCCAACAAAGCAGTCAATCAACAAATCGGGGA	1979
QY	1923	TGTATGCTGAAGAACTTCAAGGGCAGTAAAGCACTGGCGGCTCACTGTGGAAGCACAG	1982
Db	1980	CATCTTGTCTCAAGAAACATGCGCCAGTTTAAAGAGTTTACAGCTTCTTCAAAAGCATGA	2039
QY	1983	CGAGGCTCTTGAAGGCTCTGGAAGATGAATCAAGAGCTCCGCGGGCTGGAACCTTCTG	2042

Db	2040	CGAGTCTCTAACAAACTGGAAAAAGGCTACCAACGCTGTAAAGAGTTGAGGCACTGTA	2099
OY	2043	CAGAGACTTTGAGCTGCGAGAAAGGTGTATTACTACACGCTCAACACTTTCCTCTCGCGCC	2102
Db	2100	CAAGAGATTGAGCTGCGAAGAGGTCTGTCTATTGTCCTTCAACACTTCTCTGCGAAGCC	2159
OY	2103	ACTGACCCGGCTCATGCATTCACAGACGGTCTTGAGACGGCTGTGCAACACACCCGCC	2162
Db	2160	CATCCACCGCGCTGTGCATACCGCTCTGTGCGCGCGCTTATGCGGACATTACACGCC	2219
OY	2163	GAGCCACGCCGACTTCAGAGGACTGCGCAGCGCTTTTGGCAGAGATCACGAGATGTGGC	2222
Db	2220	CGGGCCACCATGACTACCGCTGACTGCAATGACGCGCTTAAAGCCATTCACAGAGGTGACAC	2279
OY	2223	ACACCTCCACGGTTCAGATGATCAAGATGAGAAATTTCCAGAAAGCTGCGACCACTCAAGAA	2282
Db	2280	CACACTACAGACATTTCTCATCCGGCTGAGAGAACCTGCABAAGCTAACGGAGCTGCGACG	2339
OY	2283	AGATTGATTTGGCATTTGACATCTTGTGTTCCGGGAGGAGAGTTTCATCCGCTGTGGCAG	2342
Db	2340	GGAAGTGTGGGCTATAGAAACCTCATTTGCTCTGGCAGAGGAGTTTCATCCGTGAGGGCTG	2399
OY	2343	CCTCAGCAAGCTCTCGGGGAGAAAGGGGCTCAGACGCGATGTTCTTCTGTTCACACGCT	2402
Db	2400	CTTTCACAAAGTCAACCAAGAGGGCTGCGACGAGAGATGTTTTTCTGTCTCAGATAT	2459
OY	2403	CCGTCTATACACAGCGCGGGGGCTGACGGGCTCCAAATCAGTTTAAAGTTCACGGGGCAGCT	2462
Db	2460	GTTCTGTGTCACAGACCAAGAGATTTGACAGGACCAAGCACTTCGGATCCGGGGCTTCCT	2519
OY	2463	CCGCGCTTATGGCATGACGATTTGAGAGAGCGCAAGACGATGGGGGGTGGCCCACTGCGCT	2522
Db	2520	TCCCTCCAAAGGCACTGCTGGTGGAAAGAAAGTATTAAGATGGCTGTTCACACTGTTT	2579
OY	2523	GACCTTCGGGGCCAGGGGCAATTCATCTGTGGCCGCAATTCGTGGTTCGAGATGGA	2582
Db	2580	CACCATTTAAGCGGCTCAGAAAACAAATCGTGTGGCAGCAGACCTCGGCTGAGAAAGA	2639
OY	2583	GAACTGGGTTGAGGACATCCAGATGGCCATTGACCTGGCGAGAAAGACGACGCCG	2642
Db	2640	GAACTGATGTGGACTGTGAATCTCCGCAATCCAGACCCAAAGATGGCCGTACACAGCG	2699
OY	2643	CCCTGAGTTCCTGGCCAGCA-----GCCGCCCTGACAAACAGTCCCTGATGAGAGCAC	2696
Db	2700	CCCTGCACTGCGACAGGCGGCACTGTGTGTCATCGTCCCCCAGATCCCCCAAGAGTATC	2759
OY	2697	CGCGGCTGACAGGAGTCAAGAGATGACTGTAGCGCTCGCGCAATTCGCTGAGCGCCA	2756
Db	2760	TCTT-----GAGAGAGAGTCAAGAAAGTATGTGTGGGGGTGTCCGAGCTCCCTGAGGGGCA	2816
OY	2757	GGCCCGGACCGCGCGCAACCAATATGGGCACTGTGTGTGGCAACGCAACCAAGCGCTC	2816
Db	2817	TGCGCAGACCGGGCCAAACACCAATGACAGTGTGTGGTACCGAAGACCAAGCGTGT	2816
OY	2817	CATGTGGAATTGACATTCGCAATGAGAAATCAGTTGTCTGAAACCTGTCTGAGAAATT	2817
Db	2877	CAGGGCAGACCAAGTGCAGCTGTGCAGAACAGCTTTCAGGATATCTGTGTAGAAAGTT	2936
OY	2877	CAAAACAGCAACGGGTGGGCAAGACTGTGGGTGTGTTCACAAATCTTCTGCTGTTCCTT	2936
Db	2937	CAAAACAGTCAATGGCTGGGCAAGAGCTCTGGGTGTGTTCACAAATCTTCTGTGTTCCTT	2996
OY	2937	CTAACAAATCAACAGGACAAATATCCCTTGGCAGCTGCGCTGTGTGGGCTATCTGCT	2996
Db	2997	CTAACAAATCTATCAGGATGATCTACCACTGGGCGAGCTTCGCGCTGTGGGCTTACAGCT	3056
OY	2997	CACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTAGTGTTCAGAGCTGCACTTCAA	3056
Db	3057	GAGCATCCCAAGGAGGGCCATGGCAATCAACAAAGACTAAGTTTTCAGAGCTTCAAGTTCAA	3116
OY	3057	GTCGCACGTCTACTTCAAGGCGGAAAGCGAGTACAGTTTCGAAAGGTGAGTGAAGT	3116

Db 3117 ATCCGAGTCTACTCTTCCGGGCTGAGAGCAAGTACATTTGAAAGGTGATGAGGT 3176
QY 3117 GATCGGAGTGGCAGCAGCTCTGCTCGGAGCCC 3151
Db 3177 GATCCAGGGGGCAGCAGCTCAGCCGGAGGGCCC 3211

RESULT 8

US-09-967-768A-104/C
; Sequence 104, Application US/09967768A
; Patent No. US2002015087A1
; GENERAL INFORMATION:
; APPLICANT: Augeneus, Meana
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967, 768A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 104
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-104

Query Match 13.8%; Score 474.4; DB 3; Length 506;
Best Local Similarity 99.6%; Pred. No. 1.1e-133;

Matches 486; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2824 GACTTCAGCATGCGAGTGAAGTCAAGTGTCTGGAACCTGCTGAGAAATTCATAAAC 2883
Db 506 GACTTCAGCATGCGAGTGAAGTCAAGTGTCTGGAACCTGCTGAGAAATTCATAAAC 447
QY 2884 AGCAACGGGTGGCAGAGCTGTGGTGGTGTTCACAACTTCTGCTGTCTTACAAA 2943
Db 446 AGCAACGGGTGGCAGAGCTGTGGTGGTGTTCACAACTTCTGCTGTCTTACAAA 387
QY 2944 TCACACAGAGCAATATCCCTTGGCAGCTGCTGTGCTGGCTACTGCTCAACATC 3003
Db 386 TCACACAGAGCAATATCCCTTGGCAGCTGCTGTGCTGGCTACTGCTCAACATC 327
QY 3004 CCTCTGAGTCCGAGAAATCCAGAAAGTACAGTGTTCAGAGCTGCACTTCAAGTCCAC 3063
Db 326 CCTCTGAGTCCGAGAAATCCAGAAAGTACAGTGTTCAGAGCTGCACTTCAAGTCCAC 267
QY 3064 GTCTACTTCTTCAAGGCGGAGAAAGCACTACGTTGAAAAGTGAATGATCCGC 3123
Db 266 GTCTACTTCTTCAAGGCGGAGAAAGCACTACGTTGAAAAGTGAATGATCCGC 207
QY 3124 AATGCCACAGAGCTCTGCTCGGACACCCCAAGTGTTCAGCAAAAGTCTTGTGTAT 3183
Db 206 AATGCCACAGAGCTCTGCTCGGACACCCCAAGTGTTCAGCAAAAGTCTTGTGTAT 147
QY 3184 TGATGGCCGAGACACTGCTTCCGAGAGTGGCTTTCTGGAAGACGTTTCTTTCTT 3243
Db 146 TGATGGCCGAGACACTGCTTCCGAGAGTGGCTTTCTGGAAGACGTTTCTTTCTT 87
QY 3244 CTGTATTAATGAAGCTGTGTAATTAACACCTGTGTGAATCAAAAATCATGGCTTCC 3303
Db 86 CTGTATTAATGAAGCTGTGTAATTAACACCTGTGTGAATCAAAAATCATGG-TTCC 28
QY 3304 AGAGCTC 3311
Db 27 AGAGCTC 20

RESULT 9

US-10-843-641A-6249/C
; Sequence 6249, Application US/10843641A
; Publication No. US2005006445A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843, 641A
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873, 367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954, 531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954, 456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962, 436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962, 832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964, 824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967, 768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968, 007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969, 347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969, 708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6249
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-6249

Query Match 13.8%; Score 474.4; DB 10; Length 506;
Best Local Similarity 99.6%; Pred. No. 1.1e-133;

Matches 486; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2824 GACTTCAGCATGCGAGTGAAGTCAAGTGTCTGGAACCTGCTGAGAAATTCATAAAC 2883
Db 506 GACTTCAGCATGCGAGTGAAGTCAAGTGTCTGGAACCTGCTGAGAAATTCATAAAC 447
QY 2884 AGCAACGGGTGGCAGAGCTGTGGTGGTGTTCACAACTTCTGCTGTCTTACAAA 2943
Db 446 AGCAACGGGTGGCAGAGCTGTGGTGGTGTTCACAACTTCTGCTGTCTTACAAA 387
QY 2944 TCACACAGAGCAATATCCCTTGGCAGCTGCTGTGCTGGCTACTGCTCAACATC 3003
Db 386 TCACACAGAGCAATATCCCTTGGCAGCTGCTGTGCTGGCTACTGCTCAACATC 327
QY 3004 CCTCTGAGTCCGAGAAATCCAGAAAGTACAGTGTTCAGAGCTGCACTTCAAGTCCAC 3063
Db 326 CCTCTGAGTCCGAGAAATCCAGAAAGTACAGTGTTCAGAGCTGCACTTCAAGTCCAC 267
QY 3064 GTCTACTTCTTCAAGGCGGAGAAAGCACTACGTTGAAAAGTGAATGATCCGC 3123
Db 266 GTCTACTTCTTCAAGGCGGAGAAAGCACTACGTTGAAAAGTGAATGATCCGC 207
QY 3124 AATGCCACAGAGCTCTGCTCGGACACCCCAAGTGTTCAGCAAAAGTCTTGTGTAT 3183
Db 206 AATGCCACAGAGCTCTGCTCGGACACCCCAAGTGTTCAGCAAAAGTCTTGTGTAT 147
QY 3184 TGATGGCCGAGACACTGCTTCCGAGAGTGGCTTTCTGGAAGACGTTTCTTTCTT 3243
Db 146 TGATGGCCGAGACACTGCTTCCGAGAGTGGCTTTCTGGAAGACGTTTCTTTCTT 87
QY 3244 CTGTATTAATGAAGCTGTGTAATTAACACCTGTGTGAATCAAAAATCATGGCTTCC 3303

Db 86 CCGTATATGAAGCTGTGTAATAATTAAACACTGTGTGAAAATCAAAAACATG--TTCCC 28
Qy 3304 AGCAGCTC 3311
Db 27 AGCAGCTC 20

RESULT 10
US-09-764-868-51
; Sequence 51, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIOR APPLICATION DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 1718
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-51

Query Match 11.7%; Score 402.4; DB 3; Length 1718;
Best Local Similarity 67.1%; Pred. No. 2,4e-111;
Matches 603; Conservative 0; Mismatches 286; Indels 9; Gaps 2;

Qy 2260 CAGAAGCTGACGAATCAAGAAAGATTGATGGCATTGACAAATCTGTGTCCGGGA 2319
Db 13 CAGAACTAACGAGCTGCAAGCGGAGCTGTGGGATAGAACTCTATTGCTCTGGAC 72
Qy 2320 AAGGAGTTTCATCCGTCTGGGCAAGCTTCAGCAAGCTTCGAGGAGGAGCTCCAGCAGCGC 2379
Db 73 AAGGAGTTTCATCCGTCTGGGCAAGCTTCAGCAAGCTTCAGCAAGGAGGAGCTCCAGCAGGAGG 132
Qy 2380 ATGTTCTTCTGTTTCAACAGAGTCTGCTATACAGAGCGGGGCTGACGGCTTCAT 2439
Db 133 ATGTTTCTGTTTCTCAATATATGTTGCTGTACAAAGAAAGAGATTGACAGGAGCAGC 192
Qy 2440 CAGTTTAAATGCACAGGAGCTCCCGCTCTATGAGCATGACGATTGAGAGAGCAGAAC 2499
Db 193 CACTTCCGATCGGGGAGCTCTTCCCTCCAGAGCATGCTGTGAAAGAAATGATTAAC 252
Qy 2500 GAGTGGGAGGAGTCCCACTGCTGACCTTCGAGGAGCAGGAGCATTCATTCGTGGCC 2559
Db 253 GAGTGGTCTGTTCAACACTGTTTCAACATCTACGCGGCTCAGAAAACATTCGTGGGCA 312
Qy 2560 GCCAGTTCTCGGTCCGAGATGAGAAAGTGGTTGAGGAATCCAGATGGCCATTGACCTG 2619
Db 313 GCCAGCATCTGGCTGAGAAAGAAAGATGAGTCTGAGACTGAATCCGCGATCCAGAGA 372
Qy 2620 GCGAGAAAGAGCAGAGCCCGCTGAGTTCTTGCCAGCA-----GCCCGCTGAC 2673
Db 373 GCCAAGAGTGGGAGTGAACAGGCGCTTGAACCTGCCAGCGGAGCAGCTGTGTGCC 432
Qy 2674 AACAAAGTCCCTGAGTGAACCAACCGGCTGACAGAGATCAGAGATGACCTGAGCGCC 2733
Db 433 CCGAGATCCCCCAAGAGATATCTCTG---GAGCAGAGATCAGAGATGATCTCGGGGT 489
Qy 2734 TGGCGCAATCGCTGAGAGGCGCAGAGCGCCGCAACCGGAGCAACATGATGAGAGTGGC 2793
Db 490 GTCCGAGCTCCCTGAGAGGAGCATGGCCAGCAGCCGGGCAACCAACCAATGACGTTGTC 549
Qy 2794 TGGCAGCCGACACACAGAGCTTCTCATGATGAGATTTCAGCATGCAAGTGAAGATCAGTTG 2853
Db 550 TGGTACCGGAACACAGAGCTGTCTCAGGAGCAACAGTGCAGCTGTCAAGAACAGCTT 609
Qy 2854 TCTGAAACCTGTGAGAAATTCAAAAACAGCAACGGGTGACAGAAAGCTGTGGGTGTG 2913

Db 610 TCAGATATCTGTAAAGAAAGTTCAAAAAAGATCATGCTGGCAGAAAGCTGTGCTCTC 669
Qy 2914 TTCAAACTTCTGCTGTTTCTTCTTCAAAATACACAGCAGCAATCATCCCTTGGCAGC 2973
Db 670 TTTCACAACTTCTGTTTCTTCTTCAAAATCAATCAGATGATACCTGAGCCAGC 729
Qy 2974 CTGCTGTGCTGAGTACTGCTCAACATCCCTCTGTGATCCGAGAAATCCAGAAAGAC 3033
Db 730 CTCCGCTGCTGAGGCTACAGGCTGAGCATCCCGAGGAGGCGGATGAGATACAAAGAC 789
Qy 3034 TACGTTTCAAGCTGACCTTCAAGTCCCAAGCTTACTTCTTCAAGGCGGAGGAGTAC 3093
Db 790 TATGTTTCAAGCTTCAAGTCCCAAGCTTACTTCTTCCGAGCTGAGAGCAAGTAC 849
Qy 3094 ACGTTCAAGAGTGGATGAGAAAGTATCCGAGTGCACAGCTGCTGCTGGAGCCCC 3151
Db 850 ACATTGAAGGTGATGAGAGTATCTCAGGAGGAGCAGCAGCTACGCGGAGGAGCC 907

RESULT 11
US-09-918-995-8794
; Sequence 8794, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8794
; LENGTH: 393
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(393)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-8794

Query Match 11.2%; Score 385.6; DB 3; Length 393;
Best Local Similarity 99.0%; Pred. No. 1.5e-106;
Matches 388; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 910 TACAGATTAACCTTGAATTCCTGATGCGCAGTGGAGATTTCTGCAAGTCTTCTGAAA 969
Db 1 TACAGATTAACCTTGAATTCCTGATGCGCAGTGGAGATTTCTGCAAGTCTTCTGAAA 60
Qy 970 ATCTGTGTTGAACATCATGCTCTTCTTGAAGTCTTCTTGAAGGCGCAAGCCCAAG 1029
Db 61 ATCTGTGTTGAACATCATGCTCTTCTTGAAGTCTTCTTGAAGGCGCAAGCCCAAG 120
Qy 1030 CCGGCTCTTGAAGCGGAGGATCATCATTTGAGTTCAAGTGTGAGTCTCAAGAGGAGTT 1089
Db 121 CCGGCTCTTGAAGCGGAGGATCATCATTTGAGTTCAAGTGTGAGTCTCAAGAGGAGTT 180
Qy 1090 CTGCACTATGTTAAAGAAAGAGCATTAAGAGTGCAGTTTGAAGAGCAGCAGCAAG 1149
Db 181 CTGCACTATGTTAAAGAAAGAGCATTAAGAGTGCAGTTTGAAGAGCAGCAGCAAG 240
Qy 1150 ATTCACTTATCCGAGAGCTTGTCTTCAAGCTTACAGAACTGAAATTCGAAAGTGTGAG 1209
Db 241 ATTCACTTATCCGAGAGCTTGTCTTCAAGCTTACAGAACTGAAATTCGAAAGTGTGAG 300
Qy 1210 CAGTTCAGCAGAGCAGCAGCTTACATTTGAGAAAGGTGCGCAATCTCCAGAGGAGCCAG 1269
Db 301 CAGTTCAGCAGAGCAGCAGCTTACATTTGAGAAAGGTGCGCAATCTCCAGAGGAGCCAG 360
Qy 1270 AGCTGCCGCGAGAGAAAGAAACGAAAGTTTC 1301

Db 361 ACCTGCCGCGAGAGAGACCGAGGTTTC 392

RESULT 12

US-09-960-352-154
Sequence 154, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byate, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21 (10298) C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 154
LENGTH: 379
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 01-LIB34-020-Q1-E1-A9
US-09-960-352-154

Query Match 8.6%; Score 297; DB 3; Length 379;
Best Local Similarity 89.7%; Pred. No. 2.1e-79;
Matches 331; Conservative 0; Mismatches 35; Indels 3; Gaps 1;

QY 1635 TGAGGCGCGAGAGAGATCCCACTGATTAAGCGTACTCATAGTAAGAGTGC 1694
DB 11 TGGGGCGCGAGAGAGATCCCACTGATTAAGCGTACTCATAGTAAGAGTGC 70
QY 1695 TACCAACGAGCGAATATCTGAAGATCTCGAAGTATCACTCGTGTTCAGAGC 1754
DB 71 CACCACTGAGAGACATATCTGAAGATCTTGAAATCATCATCTCGTGTTCAGAGC 130
QY 1755 AGTGAGCAAGAGAGAGCGCATGCGGAGACATCGAAAGTCTCATATTTCCCAATTTGA 1814
DB 131 AGTGAGCAAGAGAGATCTCATGCGCGAATCTTGAAAGTCTCATATTTCCCAATTTGA 190
QY 1815 ACCTTGCAAAATTTCTACTATTTCTCAAGAAATTTGAGCAAGATTCCTGTC 1874
DB 191 ACCTTGCAAAATTTCTACTATTTCTCAAGAAATTTGAGCAAGATTCCTGTC 250
QY 1875 GGAAGCGCGCTCAATATGCCCAATCA--GAGATTACCAAGATGCGGATGTCTATGCT 1931
DB 251 GGAAGCGCGCTCAATATGCCCAATCAAGATTTACCAAGATGCGGATGTCTATGCT 310
QY 1932 GAAGAACTTCAAGGATGAGACCTGCGGCTCACTGTGGAAGACAGGAGGCTT 1991
DB 311 GAAGAACTTCAAGGATGAGACCTGCGGCTCACTGTGGAAGACAGGAGGCTT 370
QY 1992 GGAGGCGCT 2000
DB 371 GGAGGCGCT 379

RESULT 13

US-09-918-995-34662
Sequence 34662, Application US/09918995
Patent No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hybee, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 34662
LENGTH: 455
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(455)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-34662

Query Match 8.5%; Score 293.4; DB 3; Length 455;
Best Local Similarity 99.0%; Pred. No. 2.9e-78;
Matches 306; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 3134 GCTTGCTCGGAGCCCGACGCTGTGAGCCACAAAGATCTCTGTGATTTGATGCGCG 3193
DB 2 GCTTGCTCGGAGCCCGACGCTGTGAGCCACAAAGATCTCT--TGATTTGATGCGCG 59
QY 3194 ACAACTGCTTCCGAGTGGGCTGCTTCTGGAAGAGTTCCTTCTGTATTAAT 3253
DB 60 ACACTGCTTCCGAGTGGGCTGCTTCTGGAAGAGTTCCTTCTGTATTAAT 119
QY 3254 GAAGCTGTAAATTAACACCTGTCTGAATAATCAAAACATGCTTCCAGAGCTTC 3313
DB 120 GAAGCTGTAAATTAACACCTGTCTGAATAATCAAAACATGCTTCCAGAGCTTC 179
QY 3314 CTGTCTCACAGCGCGTGTGTTTAAACCCGACCTCTCAGCGTTGATGAACAGCGCTCC 3373
DB 180 CTGTCTCACAGCGCGTGTGTTTAAACCCGACCTCTCAGCGTTGATGAACAGCGCTCC 239
QY 3374 CACTTCAGTCTGCGATCCGCTGGGGGGCGCTGTTCTTAGTAGTGCAGTATTAAAC 3433
DB 240 CACTTCAGTCTGCGATCCGCTGGGGGGCGCTGTTCTTAGTAGTGCAGTATTAAAC 299
QY 3434 ATTGTCAAT 3442
DB 300 ATTGTCAAT 308

RESULT 14

US-09-960-352-156
Sequence 156, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byate, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21 (10298) C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 156
LENGTH: 400
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 01-LIB34-020-Q1-E2-A9
US-09-960-352-156

Query Match 8.5%; Score 292; DB 3; Length 400;
Best Local Similarity 87.5%; Pred. No. 7.2e-78;
Matches 343; Conservative 0; Mismatches 45; Indels 4; Gaps 2;

QY 1635 TGAGGCGCGAGAGAGATCCCACTGATTAAGCGTACTCATAGTAAGAGTGC 1694
DB 9 TGGGGCGCGAGAGAGATCCCACTGATTAAGCGTACTCATAGTAAGAGTGC 68
QY 1695 TACCAACGAGCGAATATCTGAAGATCTCGAAGTATCACTCGTGTTCAGAGC 1754
DB 69 CACCACTGAGAGACATATCTGAAGATCTTGAAATCATCATCTCGTGTTCAGAGC 128

Search completed: May 22, 2006, 11:24:23
Job time : 3957.83 secs

```
Oy 1755 AGTGAACAAGAGAGCGCCATCGCGAAGACAGTGAAGTCTCATATTCGCCAATTTTGA 1814
    |||||
Db 129 AGTGAGCAAAAGAGAGCTCATGCCGAAACCTTGAAATCTCATATTCGGAATTTTGA 188
    |||||
Oy 1815 ACCTTTGACAAATTTTCATATTTTCTCAAGAAATTTGAGCAAGCATTTGCCCTGTG 1874
    |||||
Db 189 ACCTTTGACAAATTTTCACACATATTTTCTCAAGGACATTTGAGCAAGCATTTGCCCTGTG 248
    |||||
Oy 1875 GGAAGGCGCTGCAATGCCCAATCA--GAGATTAACCAAGATCGGAGATGCATGCT 1931
    |||||
Db 249 GGAAGGCGCTGCAATGCCCAATCAAGAGATTAACCAAGATCGGAGATGCATGCT 308
    |||||
Oy 1932 GAAGAACATTTGAGGCGATGAAGCACTTG-GCGGCTCACTGTGAAAGCAGAGGCT 1990
    |||||
Db 309 GAAGAACATTTGACGGGATGAAGCACTGAGCTGCTCATTTGTGAAGCAGAGTAAAGCCC 368
    |||||
Oy 1991 TGGAGGCGCTGGAAGATGAATCAAGAGCTCC 2022
    |||||
Db 369 TGGAGGCGCTGGAAGATGAATCAAGAGCTCC 400
    |||||
```

RESULT 15

```
US-10-062-674-58
; Sequence 58, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaeber, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/525,102
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PERL Program
; SEQ ID NO 58
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: GenBank ID No. US20040005559A1 g2305969
US-10-062-674-58
```

Query Match 8.3%; Score 286.8; DB 7; Length 437;

Best Local Similarity 88.1%; Pred. No. 3e-76; Mismatches 42; Indels 0; Gaps 0;

Matches 312; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Oy 2848 CAGTGTCTGGAACCTGCTGAGAAATCAAAAACAGCAACGGGTGCGAAGAGCTGTG 2907
    |||||
Db 1 CAGCTCTCTGGGAACCTGCTGGAAGATTCAAGAAACAGCAATGGCTGGCAAGAGCTGTG 60
    |||||
Oy 2908 GTGATGTTCAAACTTCTGCTGTTCTTCTAATAATCAACACAGAGCAATCATCCCTT 2967
    |||||
Db 61 GTGATATTCACCACTTCTGCTGTTCTTCTAATAATCAACACAGAGCAATCATCCCTT 120
    |||||
Oy 2968 GCCAGCTGCTCTGCTGCTGCTAATCGCTCACCATCCCTCTGATCCGAAACATCCAG 3027
    |||||
Db 121 GCCAGCTGCTCTGCTGCTGCTAATCGCTCACCATCCCTCTGATCCGAAACATCCAC 180
    |||||
Oy 3028 AAAAGTACGTGTTCAAGCTGACATTCAAGTCCCAAGTCTACTACTCTCAGGGCGGAAAG 3087
    |||||
Db 181 AAGGACTATGTTTCAAGCTGCAATTCAGATCCCAAGTCTACTACTCTCAGGGCTGAAGC 240
    |||||
Oy 3088 GAGTACAGTTGGAAGGTGATGAGTGAAGTATCCGAGTCCAGCAAGCTCTGCGCA 3147
    |||||
Db 241 GATATCAATTTGAAGAGGTGATGAGTGAAGTATCCGAGTCCAGCAAGCTCTGCGG 300
    |||||
Oy 3148 CCCACGTTTGAAGCAAAAGTCTCTTGTGATTTGATGCGCGGACACACTC 3201
    |||||
Db 301 GCCCAGTCTTGAATGACAAAAGATCTCACTGTACTGAGGCTGAGACAGCCC 354
    |||||
```



```

Qy 345 ACCCATTTGTGAACAGATTAGAAAGGCCAAGACAGTTGTTGTTAAAGTTTGTTGTAAGTT 404
Db 978 ABAATTAATAAAAGGAGGTTGCTGG---TGTCCCTTGGAAATTTTACATTTAATGAAGTT 1034
Qy 405 CTTTCGGCTGACACACACAACTCCAGAAGAACTACAGAGTACCTTTGGCGCTGCA 464
Db 1035 TTATCCACCTGACCCGACAGTTTAAACAGAAACATACAAAGATATTTATTTATGCTTCA 1094
Qy 465 GGTGAAGCAGGACCTGGCTCAAGGAGGTTGAAGTGTATGACACAGGCGCTCTCT 524
Db 1095 GCTTGGCGAGAACATGTTGACAGAGCTCTGCCCTGTTCTTGGACACTTGAATTAAT 1154
Qy 525 GATTTCAACATTTGTGCAATCTGAATTTGGGATTTTGAAGACCTTGAACAGAGCA 584
Db 1155 AGTTTCTTACACATCTCACTGAACTGGAGACTGACCCAGAACTCCATGGGCTGA 1214
Qy 585 CTTAGCAAAAATTAATACATACC-----TCAGCAAGACGCACTGAGAGCAAAATGCT 638
Db 1215 TTATGTTAGTGAATTTTAACTGGCCCCGAAATCAGACCAAGAACTTGAAGAGAGTCA 1274
Qy 639 GGAATTTCCCAATTAACCAATTGGAACAAACACAGAGAAATTCAGTCTCTAGA 698
Db 1275 GGAATGCTAATAGTACATACAGGTCATGACTCAGCTCAGCTGACTTGAAGTTCTTGA 1334
Qy 699 GATTGCCCCGCTGAGATGATGAAATCCGTTTGAACCCGCGCCAGAGCAGGGAAG 758
Db 1335 GAATGCCAAAAGTTGTCTATGATGAGATGATGATTTCTTAAAGCAAGAACTGGGAAG 1394
Qy 759 CACGAAGATCAATCTGGCCGCTTGCCAAACAGGAAATTCAGTGTTCAGAGGTTCACTAA 818
Db 1395 AGTAGATATCATCTTAAGGTCTGCTCTGAGGCCCTTGTGTTTACAAAGATTAAGCTGAG 1454
Qy 819 GATCAATGCTCTCAACTGAGGCCAAGGTGGGAGAGCTGAGCTTCAAGAGGAAGCGCTTCT 878
Db 1455 AATTAAACCGCTTCCCTGGCCCAAGTGTGTAAGATTTCTTAAACGTAAGTCTTTT 1514
Qy 879 CATCAAGCTCCGCGCAGATGCCAATAGTCCGTAACAGATACCTTGAATTCCTGATGGC 938
Db 1515 CATCAAGATTTGGCCCTGGAGAGCAAGAGCATATGAATGATCAATGAGATCAAACTTCC 1574
Qy 939 CAGTGGGATTTCTGAAATCTCTTCTGAAATTTCTGTGTAACATCATGCTTCTTTG 998
Db 1575 CAGTTAACGAGCAGCTTAAGAAATTAATGAAGTGTGTGAACATCACACGTTTTTCAG 1634
Qy 999 ACTTTTGAAGACCCAAACCAAGCCCAAGCCGCTCTTTAAGCCGGGGTCACTATT 1058
Db 1635 A---TTGACATCTACAGACACCATTTCCAAAAGCAATTTCTTGGCTAGATCCAAAT 1691
Qy 1059 TCGGTTCACTGGTCCGACTCAGAAAGCAG 1086
Db 1692 TCGATTCAGTGGCCGGAATCAAGCTCAG 1719

```

```

RESULT 2
US-10-473-173-56
; Sequence 56, Application US/10473173
; Publication No. US20060088823A1
; GENERAL INFORMATION:
; APPLICANT: VAN ANDEL INSTITUTE
; TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
; FILE REFERENCE: 38345-170094
; CURRENT APPLICATION NUMBER: US/10/473,173
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US 60/279,411
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 498
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 3953
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-173-56

```

```

Query Match 1.7%; Score 56.8; DB 6; Length 3953;
Best Local Similarity 46.6%; Pred. No. 5.5e-06;
Matches 218; Conservative 0; Mismatches 247; Indels 3; Gaps 1;

Qy 178 ATCCAGATCTGTGATGACACCAGAGGCAATTTGAAGTTCCACAAAGAGCTCTGGGAAG 237
Db 284 ATCCAACTGCTTAATAAGAGTTTGTGAGTTTCAACCTGCTCGTGAAGACACTGGCGCAG 343
Qy 228 GTGCTGCTGATGATGCAATTTGCAACCACTCAACCTCGTGAAGAGTGAATTTTGGCCCTC 297
Db 344 GAAAGCTTGAGGCGGTGCGCCAGAGCTGAGAGCTGCGAGAGTCACTTACTTCAGCCCTC 403
Qy 298 GAGTTTCTGATCAAAAAGATCAAGCTGTGCTGATATCTCTTAAACCCATTTGAAA 357
Db 404 TGTACTACACACAAATATCAGCGCGGTGGTGAATTTGGAABAACTTTGAAGAAAG 463
Qy 368 CAGAT--TAGAAGCCAAAGACGTTGTTGAATTTGTGTGAATTTCTTCCGCT 414
Db 464 CAGCTGATTAATATGCAATTTGAACTACCGCTATTTTGGAGTGTGTTTATGTGCT 523
Qy 415 GACCAACACAACTCCAAAGAACTCAAGATGTAACCTGCTGCGCTGAGGTGAAGCAG 474
Db 524 TCAATTTCTCAGCTGACAGAGATTAACAGATATGATTAATCTGCACTGAAGAAA 583
Qy 475 GACTTGGCTCAAGCAGGTTGACGTGTATGACACACAGCGAGCTCTTGAATTTCAAC 534
Db 584 GATATCTTGAAGAGAAATTCCTTGTACCTTAAGAAAGAAATTCAGCTACAGAGCTTA 643
Qy 535 ATTGTCAATCTGAGATTTGGGATTTTATGAAGCTTTGAGACAGAGACACTTACAAA 594
Db 644 GCTGTCAAGGAGATTTTGTGACTTTGATCATGATGAATCCAGAGACTTTTCTTCAGAAA 703
Qy 595 AATAAATCATACCTCAGCAAGCAGCACTAGAGCAAAATGCTGAA 642
Db 704 TTGCTGTGTTCTTCTGTGGAGATGTTACAGATGAAGAAATGTAATGAA 751

```

```

RESULT 3
US-10-488-619-2736
; Sequence 2736, Application US/10488619
; Publication No. US20060099578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P.C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations &
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US/10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2736
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-2736

```

```

Query Match 1.2%; Score 41.8; DB 6; Length 646;
Best Local Similarity 49.8%; Pred. No. 0.025;
Matches 106; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

Qy 2055 GCTGCAAGGATGTTTACTTAACGCTCAACCTTCTCTGCGGCACTGACCGGCT 2114
Db 73 GCTGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 132
Qy 2115 CATGACATACAGACAGGCTCTGAGAGCGCTGTGCAAAACCAACCCGCGGAGCAGCCGA 2174
Db 133 CTGCAAGTATCCCTCCAGCTGCGAGAGCTCTCAAGTACACACCCCGACAGAGGA 192
Qy 2175 CTTGAGGACTGCGGAGCGCTTTGGCAGAGATCAAGAGATGTTGGCAGACAGCTTCAAG 2234
Db 193 CTTTAAGATGTTGAAGCTGCTTGCAGCGCATGAAGATGTGGCCAGCTCATACGA 252

```


/ PRIOR APPLICATION NUMBER: US 60/577,084
/ PRIOR FILING DATE: 2004-06-04
/ NUMBER OF SEQ ID NOS: 268
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 189
/ LENGTH: 2711
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-11-145-307A-189

Query Match 1.1%; Score 36.4; DB 7; Length 2711;
Best Local Similarity 56.8%; Pred. No. 1.8;
Matches 67; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 195 CACCCAGAGGAGGATTGAGTTCCACAAGAGCTCTGGGAAGTGTGCTGATGACAT 254
DB 500 CAGCCGAGGAGCGGTCTCGATCTCCAGCAAACTGTGAGGTTCTGGCGTTTGTAGT 441
QY 255 TTGCAACCACTCAACCTGTGAGAGTACTATTTTGGCTTCGAGTTTCTGATCAC 312
DB 440 CCCCACATCTTCTCCACCACTTGGGGGTGGCCACTTGGGCTTGAAGCCCTCATCAC 383

RESULT 8
US-10-505-928-380/c
/ Sequence 380, Application US/10505928
/ Publication No. US20060088532A1
/ GENERAL INFORMATION:
/ APPLICANT: Ludwig Institute for Cancer Research et al.
/ TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
/ FILE REFERENCE: 28967/39178
/ CURRENT APPLICATION NUMBER: US/10/505,928
/ PRIOR FILING DATE: 2004-08-27
/ PRIOR APPLICATION NUMBER: US 60/363,019
/ PRIOR FILING DATE: 2002-03-07
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: PatentIn 3.2
/ SEQ ID NO 380
/ LENGTH: 1323
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-505-928-380

Query Match 1.0%; Score 35.6; DB 6; Length 1323;
Best Local Similarity 42.8%; Pred. No. 2;
Matches 179; Conservative 0; Mismatches 239; Indels 0; Gaps 0;

QY 2322 GGAATTATCCGCTCTGGGAGCCTCAGCAAGCTCTCGGGAAAGGGCTCCAGACGGCAT 2381
DB 957 GGGGCTCATGGAGCGGGTTTCGACGACTGTCTGGCAATGAAGACCAAGGGGCGAG 898
QY 2382 GTTCTTTCCTGTTCAAGAGTCTGCTATACAGAGCGGGGGCTGACGGCTCCAAATCA 2441
DB 897 CCAACACAGCTGGCCACCAACATGATCCCAAGAGCTAGCCATATCTCCACCTCGGA 838
QY 2442 GTTTAAAGTCAAGGAGCACTCCGCTCTATGSCATGACATTTAGAGAGAGCAAGCA 2501
DB 837 GTCCGGGGAGCGCTGTGGGGCGGCTCTCGCCGTGTGATGACGTGCAAGAGGTGGCCAC 778
QY 2502 GTGGGGGGTCCCGCACTGCTGACCTCCGGGGCCAGCGGAGTCCATATGTGGCCGC 2561
DB 777 GCTGACCGGTTCAGAGGAAAGGACAGCCGACCGAGAGGGCGCCCAAGATGAGAAAG 718
QY 2562 CAGTTCTCGTCCGAGATGGAAGATGGGTGAGGACATCCAGATGGCCATGACCTGGC 2621
DB 717 CAGCCCGAAAGGCGACGTCCCGGACTCGCGCCCAAGCTGAGAAACACACAGGACCCCG 658
QY 2622 GGAGAAAGACAGAGCCCGCCCTGAGTTCTGSCACAGAGCCCTGACAAACAAATGC 2681
DB 657 GATTTGCAAGGTGTAGGACACCAAGCCGACAGAGGGGCAAGGCCCAAGGCGGCGG 598
QY 2682 CCCTGATGAAGCAAGCGGGCTGACAGAGAGTCAAGAGATCACTGAGGCGCTCGGCG 2739

DB 597 CCCCCACCAAGCCCCCAGGCTGGCCCCAGGCGGCGCTGCGAGGCGACCGCCGGGCGC 540

RESULT 9
US-10-511-455-1
/ Sequence 1, Application US/10511455
/ Publication No. US20060088835A1
/ GENERAL INFORMATION:
/ APPLICANT: Pickard, Benjamin Simon
/ APPLICANT: Blackwood, Douglas
/ APPLICANT: Porteous, David
/ APPLICANT: Muir, Walter John
/ APPLICANT: Moys, Ole
/ APPLICANT: Ewald, Henrik Lykke
/ TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES
/ FILE REFERENCE: 9013.63
/ CURRENT APPLICATION NUMBER: US/10/511,455
/ PRIOR FILING DATE: 2004-10-05
/ PRIOR APPLICATION NUMBER: PCT/GB03/001543
/ PRIOR FILING DATE: 2003-04-07
/ PRIOR APPLICATION NUMBER: GB0207902.8
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: GB0207904.4
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: GB0207900.2
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: GB0207901.0
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: GB0227734.1
/ PRIOR FILING DATE: 2002-11-28
/ NUMBER OF SEQ ID NOS: 94
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 4773
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-511-455-1

Query Match 1.0%; Score 35.4; DB 6; Length 4773;
Best Local Similarity 44.5%; Pred. No. 4.8;
Matches 141; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 2459 AGCTCCCGCTTATGGCATGACGATTTAGAGAGCGAAAGCATGTGGGGGTGCCCCACT 2518
DB 2219 AACTCCAGCTGATAGAGGCTGTGATGATGACAGAAACGACTTCTGATGACCTGC 2278
QY 2519 GCTTACCTCCGGGCGCAGCGGCACTCATCATGTGGCCGCACTTCTCGTCCAGA 2578
DB 2279 ACATGACATGTCTGAGAGGACGACACACAGGCCATGCAATTCAACGCTAGCTGTGC 2338
QY 2579 TGAAGAAGTGGGTGAGGACATCCAGATGGCCATTGACCTGGCGGAGAAAGACAGACGCC 2638
DB 2339 AGCAGAAAGACGACGACGAGCGGTGGGACCAAGATACAGCCATCTTGTCCAAAC 2398
QY 2639 CCGCCCTGAGTTCTGTGCGACGAGCCCTCTGACAAAGTCCCTGATGAAGCCACCG 2698
DB 2399 AGCAGAGAGGACAGCGGTGTGGGCGGACCGAGAGGACCCGTATATGACGAGAGCT 2458
QY 2699 CGGCTGACCAAGAGATCAAGATGACTGAGGCGCTCCGCGCAATCGCTTGGAGCGCAAG 2758
DB 2459 CGAGCAAGAAACAATGCGCAGAGCCGACCGCATCTTCAACCGCTGGGCGGGCAGA 2518
QY 2759 CCGCGCACCGGGCAAC 2775
DB 2519 GGAAGCTACCTGCAAC 2535

RESULT 10
US-10-473-173-50
/ Sequence 50, Application US/10473173
/ Publication No. US20060088823A1
/ GENERAL INFORMATION:
/ APPLICANT: VAN ANDEL INSTITUTE

```

1  TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
2  TITLE OF INVENTION: Carcinoma: Prognosis and Drug target Identification
3  FILE REFERENCE: 38345-170094
4  CURRENT APPLICATION NUMBER: US/10/473,173
5  CURRENT FILING DATE: 2003-09-29
6  PRIOR APPLICATION NUMBER: US 60/279,411
7  PRIOR FILING DATE: 2001-03-29
8  NUMBER OF SEQ ID NOS: 498
9  SOFTWARE: PatentIn version 3.2
10 SEQ ID NO 50
11
12 LENGTH: 909
13
14 TYPE: DNA
15
16 ORGANISM: Homo sapiens
17
18 US-10-473-173-50

```

	Query Match	Similarity	1.0%	Score 34.4	DB 6	Length 909
	Best Local	Similarity	53.8%	Pred. No. 3.4		
	Matches	71	Conservative	0	Mismatches	61
					Indels	0
					Gaps	0
QY	1300	TCCGCGCGGAGCGCGGGTGCACCCCGAGCCTTGCGCGAGAGAGAGCCCGCGGGTAACT	1355			
Db	670	TACGGCGCGCGAAGCCGAGGCCCGGCGCGGACCGCGAAGAAAGCAGAGCGCGGGCGGAC	729			
QY	1360	AAGCAGGCGGAGCGAGCCGCTCGGCGGCCACGAGAGAGAGAGAGAGGTGTTAAGAT	1419			
Db	730	CGCAAGGCGGCGCGCGCGCGCGCGACCTCTGCGCGCAGAGAGTGGCCGCCCTTGGGCGG	789			
QY	1420	AGGACCCGACAG	1431			
Db	790	CCGACCCCGACAG	801			

```

RESULT 11
US-10-505-928-275
/ Sequence 275, Application US/10505928
/ Publication No. US2006008852A1
/ GENERAL INFORMATION:
/ APPLICANT: Ludwig Institute for Cancer Research et al
/ TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
/ FILE REFERENCE: 28967/39178
/ CURRENT APPLICATION NUMBER: US/10/505,928
/ CURRENT FILING DATE: 2004-08-27
/ PRIOR APPLICATION NUMBER: US 60/363,019
/ PRIOR FILING DATE: 2002-03-07
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: PatentIn 3.2
/ SEQ ID NO 275
/ LENGTH: 4450
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-505-928-275

```

Query Match	Best Local Similarity	1.0%: Score 34.2; DB 6, Length 4450;
Matches	87; Conservative 0; Mismatches 88; Indels 0; Gaps 0;	
Qy	2106 GCACCGGCTCATGCACTAACAAGAGCTCTGAGCGGCTGTGCAAAACCAACCCGCGAG	2165
Db	2010 GCACCTGGCACAAGAAGTACTGTGGGTACAAGCCCTGGAAGCCCTCGGCTCAAGCAAA	2065
Qy	2166 CCACGCCGACTTTCAGAGGACTGCGCAAGCCGCTTTGGCAGAGATCACGAGATGTGTGCACA	2225
Db	2070 CTTGACCGACCTCTCGGTGAACGTGACGCACTCGCTGAGATGACAGTCTTTGTGGCCGG	2129
Qy	2226 GCTCCAGCGATCGATGATCAAGATGAGAAATTTCCAGAAAGTGCACGAATCTCAAG	2280
Db	2130 AGCGCAGCGCCCAAGCATCGTGTGTGTAACAAAGACAGAGGCGCTGTCTGGAGGAAAG	2184

RESULT 12
US-11-312-958-25
; Sequence 25, Application US/11312958
; Publication No. US20060100152A1

```

GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Rosenfeld, Julie Beth
APPLICANT: Sileo-Santiago, Inmaculada
TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
TITLE OF INVENTION: 12216, 41897, 47174, 33408, 10002, 16209, 314, 636,
TITLE OF INVENTION: 27410, 33260, 619, 15985, 65112, 2158, 224, 615, 44373.
TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55034, 16314, 1613, 1675, 9569 OR
TITLE OF INVENTION: 13424 MOLECULES
FILE REFERENCE: MP102-027P1R0NANIM
CURRENT APPLICATION NUMBER: US/11/312,958
CURRENT FILING DATE: 2005-12-20
PRIOR APPLICATION NUMBER: US/10/369,022
PRIOR FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: US 60/360,495
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/370,121
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US 60/373,010
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/373,908
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/377,717
PRIOR FILING DATE: 2002-05-03
PRIOR APPLICATION NUMBER: US 60/379,949
PRIOR FILING DATE: 2002-05-13
PRIOR APPLICATION NUMBER: US 60/382,409
PRIOR FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: US 60/385,280
PRIOR FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US 60/386,879
PRIOR FILING DATE: 2002-06-06
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 4234
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (863)...(2452)
US-11-312-958-25

```

Query Match	1.0%;	Score 34;	DB 7;	Length 4242;
Best Local Similarity	53.8%;	Pred. No. 11;		
Matches 70;	Conservative 0;	Mismatches 60;	Indels 0;	Gaps 0;
QY	2288	TGATTGCATTTGCAACATCTTGTGTGCTTCGGGAGAGGAGTTTCATCCGTCGGGACGCTCA	2347	
Db	1908	TGTTCCGCATCTTCAAGCTCTCCGCCACATCCAGAGGGGCTGCAAGATCCTGGGCAAGACT	1967	
QY	2348	GCAAGCTCTCGGGGAGAGGGGCTCCAGACGGCGCATGTTCTTCTGTTCACGACGTCCTGC	2407	
Db	1968	TGCGAGGCTTCATAGGAGAGGAGCTGGGAGCTGCTCATCTTCTCTTCATCGGGGTCATCC	2027	
QY	2408	TATACACGAG	2417	
Db	2028	TCTTCTCAG	2037	

```

RESULT 13
US-10-505-928-409
: Sequence 409, Application US/10505928
: Publication No. US2006008852A1
: GENERAL INFORMATION:
: APPLICANT: Ludwig Institute for Cancer Research et al
: TITLE OF INVENTION: LYMPHATIC ENDOTHelial GENES
: FILE REFERENCE: 28567/39178
: CURRENT APPLICATION NUMBER: US/10/505,928
: CURRENT FILING DATE: 2004-08-27
: PRIORITY APPLICATION NUMBER: US 60/363,019

```

```

; PRIOR FILING DATE:2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patentn 3.2
; SEQ ID NO 4079
; LENGTH: 4573
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-505-828-409

```

Query Match	1.0%;	Score 34;	DB 6;	Length 4573;
Best Local Similarity	59.2%;	Pred. No. 11;		
Matches 58;	Conservative 0;	Mismatches 40;	Indels 0;	Gaps 0;

Oy 1542 GGTCTCCCAACTGAGCCCCGACACCAAGCAGAGGCTCTCCCTTATACAGCCCGCTGCTGA 1601
 Db 437 GTGCGGATTTGGAGACAGAGCTTCCAGAGATTCTGCCCCACCATCTCTCCAGCAGACTGA 496

QY 1602 TGACCAGGCCTGCCCCGGACGACGATGAGATGAGG 1635
 Db 497 TTCCCGGGCCTGCACCTCGGAGAACGAAAAAGAGG 534

```

RESULT 14
US-10-505-928-47
/ Sequence 47, Application US/10505928
/ Publication NO. US20060088532A1
/ GENERAL INFORMATION:
/ APPLICANT: Ludwig Institute for Cancer Research
/ TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
/ FILE REFERENCE: 28567/39178
/ CURRENT APPLICATION NUMBER: US/10/505, 928
/ PRIOR FILING DATE: 2004-08-27
/ PRIOR APPLICATION NUMBER: US 60/363,019
/ PRIOR FILING DATE: 2002-03-07
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: PatentIn 3.2
/ SEQ ID NO 47
/ LENGTH: 4585
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: KIAA0062
/ US-10-505-928-47

```

Query Match	1.0%;	Score 34;	DB 6;	Length 4585;
Best Local Similarity	59.2%;	Pred. No. 11;		
Matches 58;	Conservative 0;	Mismatches 40;	Indels 0;	Gaps 0;

QY 1512 GGTCCCAACTGAGCCCGACACCAAGAGGCTCTCCCTTGATACGCCCGCTGTA 1601

D5 449 GTGCGGATTGGAGCAGGAGCTCAAGAGTTCTGCCCAACCAATCCTCCAGCAGCTGA 508

QY 1602 TGACCAAGCCTGCCCCCGGACGACGATGAGGATGAGG 1633

Db 509 TTCCCGGGCCTGCACCTCGGAGAACGAGAAACGAGG 546

RESULT 15
US-11-217-529-174434
Sequence 174434, Application US/11217529
Publication No. US2006099612A1
GENERAL INFORMATION:
APPLICANT: SUNTORY LIMITED
APPLICANT: NAKAO, YOSHITIRO
APPLICANT: NAKAMURA, NORIHISA
APPLICANT: KODAMA, YUKIKO
APPLICANT: FUJIMURA, TOMOKO
APPLICANT: ASHIKARI, TOSHITAKO
TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
FILE REFERENCE: S-38-285
CURRENT APPLICATION NUMBER: US/11/217,529
CURRENT FILING DATE: 2005-09-02
PRIOR APPLICATION NUMBER: US 10/932,182

```

; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 174434
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
; US-11-217-529-174434

```

Query Match	1.0%;	Score 33.8;	DB 7;	length 555;
Best Local Similarity	48.2%;	Pred. No. 3.7;		
Matches 95; Conservative	0;	Mismatches 102;	Indels 0;	Gaps 0;

oy 2580 GGGAGAGGGGTTGAGGACATCCAGATGGCCATTGACCTGGCCGAGAAAGAGCAGACGCC 2639
 Db 246 GCACCACTGTTTCTGGAACAAGAAACGCCATAGATCTTGAGACGACGACCCCGCACAC 305

Oy 2640 CCGCCCTGAGTTCCTGGCGAGCAGCCCCCTGTCAACAAAGTCCCTGTATGAAGCACCCTC 2659
 Db 306 AAACACTTCGATCAATATAATTAAACAACATGACAGCAGCCGCGCTTGACGAGGACGCCGA 365

2700 GGGTACACGAGATCAGAGATGACCTCGAGCGCTCGGCACATCGCTGAGCGCCACAGC 2755
 366 CCAAGACATTATTTATCAAGAGCACTCCATACAAATTCTCGTGGCCGTGATTTCTGCC 425

QY	2760	CCCCACCGCGGCAACA	2776
Db	426	CAGCAGCAGTAGCAACA	442

Search completed: May 22, 2006, 11:25:34
Job time : 36.587 secs


```
Db 392 GTTGCTCTGACACGATATTCATACGTAACTTGAGAGATGATATTTTGGGCTA 451
Qy 298 GAGTTCCTGATACAAAAAGATCAGGCTGCTGATCTCTAAACCATTGGAAA 357
Db 452 CGTTACTGTGACGAAGCCATCAGATATTTGGCTGATCTCTCAAAAACCTTGCTGAA 511
Qy 358 CA-----GATTGAAGGCCAAGCAGCTGTTGTTAAGTTTGAGTGAATCTTT 408
Db 512 CACAAAGAACTGATCAACACTGAGCTCCATATACCTTTGATTTTGGATTAATCTAT 571
Qy 409 CCGCTGACACACACACACTCCAAGAGAACTCAAGAGTCTGTTCCGCTGAGAGTG 468
Db 572 GGTGAAGATCATGTAACCTTAAGAAATAACAGATATCAGTTTTTCTTGCAAGTG 631
Qy 469 AAGCAGACTTGCTCAAGCAGCTTGACGTATATGACCAAGCAGCTCTCTTGANT 528
Db 632 AAGCAAGATGTCCTTCAAGGCGCTGCTGCTGCTCAACCTGCTCAAGCTGAG 691
Qy 529 TCACACATTGTGCACTGAGATTTGGGANTTTG-----ATGAGCCTTGACAGAG 582
Db 692 GGTATGCCATCCAGTCGAGCTTGAGATTAATGACCATATTAACACTGAGAGAT 751
Qy 583 CACTTAGCAAAAATTAATACATACCTCAAGCAGACGCACTAGAGGACAAATCGTGAA 642
Db 752 GATCTGAGTACCGGTTTGTCTGATCAGAGAGAACTTGAAGAGCATTAGAAAAG 811
Qy 643 TTTCACCATTAACCACTTGGACAAACACAGCAAACTCAGATTTCCAGCTCTAGAGAT 702
Db 812 ATTCAATAAACTTAATGGGTGAGATTTCTTCTGAGGCTGAGCTGATTAATTTGAGGACT 871
Qy 703 GCCCGTGGCTGAGATGATGATGATCCGTTGCAACCGGCGCAAGGAGGAGGACG 762
Db 872 GCAAAATCCCTGAGAGATGATGAGCTTACCTCCATGCTGAGAGAAACAAAGTCT 931
Qy 763 AAGATCATCTGGCCCTTGCCAAACAGGAAATCTAGTGTTCAGGCTTCACTAGATC 822
Db 932 GAGTATTTCTTGAATTAATCTCGGTTGTTGTTGTGATCAAGAAATAAAGCAAGTG 991
Qy 823 AATGCTTCACTGGGCGCAAGGTGCGAAGCTGAGCTTCAAGAGAAAGCGCTTTCATC 882
Db 992 GGAAGATATTTCTGGCTCGGATTAACAAAGTTCACCTCAAGAGACTCAATTTGAATC 1051
Qy 883 AAGCTCCGCGCAGATCCAAATAGTGGTACAGAGATACCTTGAATTCCTGAGGCCAGT 942
Db 1052 AAGATCTGAGAAAGATGTAACGAACTCATTTTGAAGCTGG-----AGT 1105
Qy 943 CCGGATTTCTGCAAGTCTTCTGAAAATCTGTGTTGAACATGCTCTTTAGACTT 1002
Db 1106 AAAACTGCTTGCAAGACCTCTGGAAGTGAGTGGAACATCATATTTTAAAGATG 1165
Qy 1003 TTTGAAGA 1010
Db 1166 CCAGAAAA 1173

RESULT 4
US-11-418-718-2805
; Sequence 2805, Application US/11418718
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benitach, Itzhak
; APPLICANT: Avniel, Amir
; APPLICANT: Karov, Yael
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN LIVER LUNG AND BLUDDER CANCER
; FILE REFERENCE: Efs-56
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 8661
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2805
; LENGTH: 63
```

```
; TYPE: RNA
; ORGANISM: Human
US-11-418-718-2805
Query Match 1.6%; Score 56; DB 7; Length 63;
Best Local Similarity 67.9%; Pred. No. 1e-05;
Matches 38; Conservative 18; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3387 GGCATCCGCTGGGGGCGCTGTTCTTTAGCTAGTCCAGATTAATAACATTGCATT 3442
Db 1 GGCATCCGCTGGGGGCGCTGTTCTTTAGCTAGTCCAGATTAATAACATTGCATT 3442

RESULT 5
US-11-418-870-2805
; Sequence 2805, Application US/11418710
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benitach, Itzhak
; APPLICANT: Avniel, Amir
; APPLICANT: Karov, Yael
; APPLICANT: Aharonov, Ranit
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN LIVER LUNG AND BLUDDER CANCER
; FILE REFERENCE: Efs-56
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 8661
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2805
; LENGTH: 63
; TYPE: RNA
; ORGANISM: Human
US-11-418-870-2805
```

```
Query Match 1.6%; Score 56; DB 7; Length 63;
Best Local Similarity 67.9%; Pred. No. 1e-05;
Matches 38; Conservative 18; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3387 GGCATCCGCTGGGGGCGCTGTTCTTTAGCTAGTCCAGATTAATAACATTGCATT 3442
Db 1 GGCATCCGCTGGGGGCGCTGTTCTTTAGCTAGTCCAGATTAATAACATTGCATT 3442
```

```
RESULT 6
US-11-418-875-2805
; Sequence 2805, Application US/11418715
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benitach, Itzhak
; APPLICANT: Avniel, Amir
; APPLICANT: Karov, Yael
; APPLICANT: Aharonov, Ranit
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN LIVER LUNG AND BLUDDER CANCER
; FILE REFERENCE: Efs-56
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 8661
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2805
; LENGTH: 63
; TYPE: RNA
; ORGANISM: Human
US-11-418-875-2805
```

```
Query Match 1.6%; Score 56; DB 7; Length 63;
Best Local Similarity 67.9%; Pred. No. 1e-05;
Matches 38; Conservative 18; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3387 GGCATCCGCTGGGGGCGCTGTTCTTTAGCTAGTCCAGATTAATAACATTGCATT 3442
Db 1 GGCATCCGCTGGGGGCGCTGTTCTTTAGCTAGTCCAGATTAATAACATTGCATT 3442
```

```
RESULT 7
US-11-418-718-5577
; Sequence 5577, Application US/11418718
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benitwch, Itzhak
; APPLICANT: Avniel, Amir
; APPLICANT: Karov, Yael
; APPLICANT: Aharonov, Rantl
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN LIVER LUNG AND BLUDDER CANCER
; FILE REFERENCE: Efs-56
; CURRENT APPLICATION NUMBER: US/11/418,718
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 8661
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5577
; LENGTH: 62
; TYPE: RNA
; ORGANISM: Human
US-11-418-718-5577

Query Match      1.6%; Score 55; DB 7; Length 62;
Best Local Similarity 67.3%; Pred. No. 1.9e-05;
Matches 37; Conservative 18; Mismatches 0; Indels 0; Gaps 0;

OY      3388 GCATCCGCTGGGGGCGCTGTTCTTAGCTAGTGCAGATTAATAACATTGTCATT 3442
DB      1 GCAUCCGCGUGGGGCGCGUCUUCUUAAGCUAGUCCAGUAUUAAAACAUUGUCAU 55

RESULT 8
US-11-418-870-5577
; Sequence 5577, Application US/11418870
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benitwch, Itzhak
; APPLICANT: Avniel, Amir
; APPLICANT: Karov, Yael
; APPLICANT: Aharonov, Rantl
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN LIVER LUNG AND BLUDDER CANCER
; FILE REFERENCE: Efs-56
; CURRENT APPLICATION NUMBER: US/11/418,870
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 8661
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5577
; LENGTH: 62
; TYPE: RNA
; ORGANISM: Human
US-11-418-870-5577

Query Match      1.6%; Score 55; DB 7; Length 62;
Best Local Similarity 67.3%; Pred. No. 1.9e-05;
Matches 37; Conservative 18; Mismatches 0; Indels 0; Gaps 0;

OY      3388 GCATCCGCTGGGGGCGCTGTTCTTAGCTAGTGCAGATTAATAACATTGTCATT 3442
DB      1 GCAUCCGCGUGGGGCGCGUCUUCUUAAGCUAGUCCAGUAUUAAAACAUUGUCAU 55

RESULT 9
US-11-429-720-2923
; Sequence 2923, Application US/11429720
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benitwch, Itzhak
; APPLICANT: Avniel, Amir
; APPLICANT: Karov, Yael
; APPLICANT: Aharonov, Rantl
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN PROSTATE CANCER
; FILE REFERENCE: Efs-59
; CURRENT APPLICATION NUMBER: US/11/429,720
; CURRENT FILING DATE: 2006-05-08
```

```
; NUMBER OF SEQ ID NOS: 3751
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2923
; LENGTH: 62
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-429-720-2923

Query Match      1.6%; Score 55; DB 7; Length 62;
Best Local Similarity 67.3%; Pred. No. 1.9e-05;
Matches 37; Conservative 18; Mismatches 0; Indels 0; Gaps 0;

OY      3388 GCATCCGCTGGGGGCGCTGTTCTTAGCTAGTGCAGATTAATAACATTGTCATT 3442
DB      1 GCAUCCGCGUGGGGCGCGUCUUCUUAAGCUAGUCCAGUAUUAAAACAUUGUCAU 55

RESULT 10
US-11-418-875-5577
; Sequence 5577, Application US/11418875
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benitwch, Itzhak
; APPLICANT: Avniel, Amir
; APPLICANT: Karov, Yael
; APPLICANT: Aharonov, Rantl
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN LIVER LUNG AND BLUDDER CANCER
; FILE REFERENCE: Efs-56
; CURRENT APPLICATION NUMBER: US/11/418,875
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 8661
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5577
; LENGTH: 62
; TYPE: RNA
; ORGANISM: Human
US-11-418-875-5577

Query Match      1.6%; Score 55; DB 7; Length 62;
Best Local Similarity 67.3%; Pred. No. 1.9e-05;
Matches 37; Conservative 18; Mismatches 0; Indels 0; Gaps 0;

OY      3388 GCATCCGCTGGGGGCGCTGTTCTTAGCTAGTGCAGATTAATAACATTGTCATT 3442
DB      1 GCAUCCGCGUGGGGCGCGUCUUCUUAAGCUAGUCCAGUAUUAAAACAUUGUCAU 55

RESULT 11
US-11-431-708-6398
; Sequence 6398, Application US/11431708
; GENERAL INFORMATION:
; APPLICANT: RUBEN, Steven
; TITLE OF INVENTION: BREAST DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CLO01599-ORD
; CURRENT APPLICATION NUMBER: US/11/431,708
; CURRENT FILING DATE: 2006-05-11
; NUMBER OF SEQ ID NOS: 8544
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6398
; LENGTH: 5857
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-431-708-6398

Query Match      1.2%; Score 40.4; DB 7; Length 5857;
Best Local Similarity 46.1%; Pred. No. 0.66;
Matches 177; Conservative 0; Mismatches 201; Indels 6; Gaps 1;

OY      2059 CAGAAGTGTGTTACTTACCGCTCAACACCTTCCTCGGCGCACTGCACGCGCTCATG 2118
DB      564 CGAAGACACACGAGACATCCCTTTGGAAGGCTTACTCTGTTGTCTCCGATCCAGAGATCTCG 623
OY      2119 CACTACAAGCAGGTCCTGAGGCGCTGTGCAAAACACCAACCGCGAGCCAGCGGACTTC 2178
```

Db 624 AGTACCCGCTCTCTTAAAGAGCTGGCCAGAGAGACTCCGGCAGAGACCCAGAC 683
Qy 2179 AGGAGCTCCGAGCCGCTTTGGCAGAGATCACGAGATGGTGCCACAGCTCCACGGTACG 2238
Db 684 CCGCGGTCAGAGTCCCTGCGAGGCCATGAGACCGTTTGCTCCAACATCATATGAGACC 743
Qy 2239 ATGATCAAGATGAGAAATTTCCAGAACTGCAGAACTCAAGAAATTTGATGGCATT 2298
Db 744 AAGCGGCAATGAGAACTGGAAAGCCCTGAGACACTGACATGCCAATCGAAGGCTGG 803
Qy 2299 GA-----CAATCTTGTGTTCCGGGAAGGAGTTATCCGTCGTGGCAGCCTCAGCAAG 2352
Db 804 GAGGGTCCAACTCACAAGACATCTGACCTCTCTGCAAGGACTTTGTTAAAG 863
Qy 2353 CTCTCGGGAGAGGGCTCCAGACAGCCAGTCTTCTGTTCAACAAGTCTGCTATAC 2412
Db 864 ATCTCTGCGGGCAACATCCAGGAAAGGGCTTCTCTTCCAGCAACCTTCTGCTAC 923
Qy 2413 ACGAGCCGGGGGCTGACGGCTCC 2436
Db 924 TGCAGCGGAATCCAGGGTCAAC 947

RESULT 12
US-10-669-920-1193
/ Sequence 1193, Application US/10669920
/ GENERAL INFORMATION:
/ APPLICANT: Morris, David W.
/ APPLICANT: Malandro, Marc S.
/ TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
/ FILE REFERENCE: 20366-066001
/ CURRENT APPLICATION NUMBER: US/10/669,920
/ CURRENT FILING DATE: 2003-09-23
/ PRIOR APPLICATION NUMBER: US 10/004,113
/ PRIOR FILING DATE: 2001-10-23
/ PRIOR APPLICATION NUMBER: US 10/052,482
/ PRIOR FILING DATE: 2001-11-08
/ PRIOR APPLICATION NUMBER: US 09/997,722
/ PRIOR FILING DATE: 2001-11-30
/ PRIOR APPLICATION NUMBER: US 10/034,650
/ PRIOR FILING DATE: 2001-12-20
/ PRIOR APPLICATION NUMBER: US 10/085,117
/ PRIOR FILING DATE: 2002-02-27
/ PRIOR APPLICATION NUMBER: US 10/087,192
/ PRIOR FILING DATE: 2002-03-01
/ PRIOR APPLICATION NUMBER: US 10/322,281
/ PRIOR FILING DATE: 2002-12-17
/ PRIOR APPLICATION NUMBER: US 10/322,696
/ PRIOR FILING DATE: 2002-12-17
/ NUMBER OF SEQ ID NOS: 1441
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1193
/ LENGTH: 63810
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(63810)
/ OTHER INFORMATION: n = A,T,C or G
US-10-669-920-1193

Query Match 1.1%, Score 38.8, DB 6, Length 63810,
Best Local Similarity 49.0%, Pred. No. 4.5,
Matches 103, Conservative 0, Mismatches 107, Indels 0, Gaps 0,

Qy 2460 GCTCCGCTCTATGATGACATGAGAGAGCAAGATGGGGGTGCCCACTG 2519
Db 30746 GCGCCAGGTCGTGTATGTAAGTCTGACCGGAGAGAGCCGTGAAGACCCCGCT 30805
Qy 2520 CCTGACCTTCGGGGCCAGCGGAGTCCATCATCTGGCGGCCAGTTCTCGTCCAGAT 2579
Db 30806 GCTGTCTTACGTCGACACTGAAGGCCGAGCAAGGGCTGACCATGTGGGCTCGTGTCT 30865

Qy 2580 GGAGAGTGGGTTGAGACATCCAGTTCATTGACTTGGCGGAGAGAGAGAGCCCC 2639
Db 30866 GAGGGGAGCTACCTTGACAGACATGAGAGCTCAGCGGCCAGAGAGTGGGCGGGC 30925
Qy 2640 CCGCCCTGAGTTCCTGGCCAGAGCCCCC 2669
Db 30926 GGGGTCTGGGGCCAGGCTCTTTCCAC 30955

RESULT 13
US-10-461-194A-2
/ Sequence 2, Application US/10461194A
/ GENERAL INFORMATION:
/ APPLICANT: Hutchinson, Richard C.
/ APPLICANT: Reid, Ralph C.
/ APPLICANT: Hu, Zhihao
/ APPLICANT: Rascher, Andreas
/ APPLICANT: Schirmer, Andreas
/ APPLICANT: McDaniel, Robert
/ TITLE OF INVENTION: RECOMBINANT POLYNUCLEOTIDES ENCODING
/ TITLE OF INVENTION: PRO-GLUTAMINICIN PRODUCING POLYPEPTIDE SYNTHASES AND
/ FILE REFERENCE: 300622009700
/ CURRENT APPLICATION NUMBER: US/10/461,194A
/ CURRENT FILING DATE: 2003-06-13
/ PRIOR APPLICATION NUMBER: US 60/389,255
/ PRIOR FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: US 60/393,929
/ PRIOR FILING DATE: 2002-07-03
/ PRIOR APPLICATION NUMBER: US 60/395,275
/ PRIOR FILING DATE: 2002-07-12
/ PRIOR APPLICATION NUMBER: US 10/212,962
/ PRIOR FILING DATE: 2002-08-05
/ PRIOR APPLICATION NUMBER: US 60/415,326
/ PRIOR FILING DATE: 2002-09-30
/ PRIOR APPLICATION NUMBER: US 60/420,820
/ PRIOR FILING DATE: 2002-10-24
/ PRIOR APPLICATION NUMBER: US 60/433,130
/ PRIOR FILING DATE: 2002-12-13
/ NUMBER OF SEQ ID NOS: 161
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 86941
/ TYPE: DNA
/ ORGANISM: Streptomyces hygroscopicus
US-10-461-194A-2

Query Match 1.1%, Score 38.6, DB 6, Length 86941,
Best Local Similarity 42.0%, Pred. No. 5.8, DB 309, Indels 0, Gaps 0,
Matches 224, Conservative 0, Mismatches 309, Indels 0, Gaps 0,

Qy 2322 GAGATTATCCGCTGCGGAGCTTCAGCAAGCTTCGGGGAAGGGCTCCAGAGGCAT 2381
Db 61950 GAGGGCCGAACGGTTCGATCAACCGCGCTGCTGAGACGCCGCTCAGCGGGCAG 62009
Qy 2382 GTTCTTCTGTTCAAGAGTCTGTATACAGAGCCGGGGCTGACGCTCCATATA 2441
Db 62010 CTTCGCTGCTGCTGAGCCGGCGGAGACAGTGAATCTGCTGCTGCTGGAACATC 62069
Qy 2442 GTTTAAAGTCAAGGAGCACTCCCGCTATGAGCATGACATTTAGAGAGAGCAAGACA 2501
Db 62070 CGTGGCTGACACCGCGGGGCGGCGCTCGCGGTCCGGGTGATGTCGCCCGGCGCA 62129
Qy 2502 GTGGGGGAGTCCCACTGCTGACCTTCGGGGCCAGCGGCACTCATATGTCGCGC 2561
Db 62130 CAGCGCTTCTGATGATGCTGACCGAGCGCTCGGGCCAGAGGTGGCTCCGTGGAATC 62189
Qy 2562 CAGTTCTCGGTCCGAGATGAGAAATGGTTAGACATCCAGATGCGCATGACCTGAGC 2621
Db 62190 GCTCACTTGGCGCGGAGTGAATCCGCGCACTGAGATGAGTGAAGCCGAGAGCGGCT 62249
Qy 2622 GAGAGAGACAGAGCCCGCCCTGAGTTCTGTGCGCAAGAGCCCGCTGACAAAGTC 2681

Db 62250 GTGACGGTCCGTTGGAGGAGACCTCCCTGCCGAGCGCGGATCTCTGGAGCGCGCT 62309
Qy 2682 CCTGATGAAGCACCGCGCGCTGACCAAGAGTCAGAGTACCTGAGGCGCTCGGCGAC 2741
Db 62310 GGGCGATCCGGCCACCGGTGGCGCGCGGCGCATGGAGACCGGAAGCGCGGAGCGCG 62369
Qy 2742 ATCGCTGAGCGCCAGCGCGCGCGCGCAACAATGTGTGACGTGTGCTGGCAGCG 2801
Db 62370 CCTCCCGACGTCCTCTGTGGCGATACGGCGCGCTGGCGCGCAAGACCTCACCGGCGCGC 62429
Qy 2802 CAACACAGCGTCTTCATGATGTGACCTTCAGCATTCGATGGAGATCACTGT 2854
Db 62430 GACCGCGCGCGCGCGCGCGCTCACCGCGCGCTGTGGCGGAGATCAAGCGGT 62482

RESULT 14
US-10-669-920-398/c
; Sequence 398, Application US/10669920
; GENERAL INFORMATION:
; APPLICANT: Moris, David W.
; APPLICANT: Malandro, Marc S.
; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
; FILE REFERENCE: 20366-066001
; CURRENT APPLICATION NUMBER: US/10/669,920
; PRIOR FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: US 10/004,113
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 10/052,482
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 09/997,722
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 10/034,650
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 10/085,117
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 10/322,281
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 10/322,696
; PRIOR FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 1441
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 398
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-669-920-398

Query Match 1.1%; Score 37.2; DB 6; Length 1394;
Best Local Similarity 43.1%; Pred. No. 2.4;
Matches 180; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
Qy 2322 GGAGTTCATCCGTCGGGAGCCTCAGCAAGCTTCGGGAAAGGGCTTCAGACGGCAT 2381
Db 1120 GGGGCTCATGGAGGAGGGGTTTCGACGACTGTCTGGGAGATGAAGACAGAAAGGGCAG 1061
Qy 2382 GTTCTTCTGTTCAAGAGAGCTCTGCTATACAGAGAGCCGGGGCTGACGGCTCCAAATA 2441
Db 1060 CCACACAGCGTGGCGCACCATGATCCCGAGAGCTAGCCATCATCTTCCACCTCGGA 1001
Qy 2442 GTTAAAGTCCAGCGGAGAGCTCCCGCTCTATGGCATGAGATTGAGAGAGGAGAGCA 2501
Db 1000 GTCCCGGGAGACGCTGTGGGCGCGCTCTGCGCGTGTAGAGCTGACAGGGGTGGCAC 941
Qy 2502 GTGGGGGTGCCCCCATCTGACCTCCGGGGCCAGCGGCACTCATCATCTGTGGCGCG 2561
Db 940 GGTGACCGGTTCAGAGAGAGAGAGAGAGCGCCGAGCGAGAGCGCGCCAGCATGAGAAAG 881
Qy 2562 CAGTTCTCGGTCCGAGATGAGAAAGTGGTTGAGAGCATTCAGATGGCCATTGACCTGGC 2621
Db 2562 CAGTTCTCGGTCCGAGATGAGAAAGTGGTTGAGAGCATTCAGATGGCCATTGACCTGGC 2621
Qy 880 CAGCCCGAAGGCGCAGTCCCGGAGCTCGCGCCAGCGCTCAGAGAACACAGAGACCCCGC 821

Qy 2622 GGAGAAGAGCAGACCGCGCGCTGAGTTCCTGGCCAGACGCGCGCTGACAAAGTTC 2681
Db 820 GTATTGACAGGTGTAGCCAGCCAGCGGCGAGAGGGCCAGAGGCCAAGGCCAGCGCGCG 761
Qy 2682 CCTGATGAAGCACCGCGCGCTGACCAAGAGTCAAGAGATGACCTGAGCGCGCTCGCGC 2739
Db 760 CGCCACACAGC 703

RESULT 15
US-10-669-920-400/c
; Sequence 400, Application US/10669920
; GENERAL INFORMATION:
; APPLICANT: Moris, David W.
; APPLICANT: Malandro, Marc S.
; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
; FILE REFERENCE: 20366-066001
; CURRENT APPLICATION NUMBER: US/10/669,920
; PRIOR FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: US 10/004,113
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 10/052,482
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 09/997,722
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 10/034,650
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 10/085,117
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 10/322,281
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 10/322,696
; PRIOR FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 1441
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 400
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-669-920-400

Query Match 1.1%; Score 37.2; DB 6; Length 1507;
Best Local Similarity 43.1%; Pred. No. 2.5;
Matches 180; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
Qy 2322 GGAGTTCATCCGTCGGGAGCCTCAGCAAGCTTCGGGAAAGGGCTTCAGACGGCAT 2381
Db 1120 GGGGCTCATGGAGGAGGGGTTTCGACGACTGTCTGGGAGATGAAGACAGAAAGGGCAG 1061
Qy 2382 GTTCTTCTGTTCAAGAGAGCTCTGCTATACAGAGAGCCGGGGCTGACGGCTCCAAATA 2441
Db 1060 CCACACAGCGTGGCGCACCATGATCCCGAGAGCTGAGCCATCATCTTCCACCTCGGA 1001
Qy 2442 GTTAAAGTCCAGCGGAGAGCTCCCGCTCTATGGCATGAGATTGAGAGAGGAGAGCA 2501
Db 1000 GTCCCGGGAGACGCTGTGGGCGCGCTCTGCGCGTGTAGAGCTGACAGGGGTGGCAC 941
Qy 2502 GTGGGGGTGCCCCCATCTGACCTCCGGGGCCAGCGGCACTCATCATCTGTGGCGCG 2561
Db 940 GGTGACCGGTTCAGAGAGAGAGAGAGAGCGCCGAGCGAGAGCGCGCCAGCATGAGAAAG 881
Qy 2562 CAGTTCTCGGTCCGAGATGAGAAAGTGGTTGAGAGCATTCAGATGGCCATTGACCTGGC 2621
Db 880 CAGCCCGAAGGCGCAGTCCCGGAGCTCGCGCCAGCGCTCAGAGAAAGACAGAGACCCCGC 821
Qy 2622 GGAAGAAGCAGACCGCGCGCTGAGTTCCTGGCCAGACGCGCGCTGACAAAGTTC 2681
Db 820 GTATTGACAGGTGTAGCCAGCCAGCGGCGAGAGGGCCAGAGGCCAAGGCCAGCGCGCG 761
Qy 2682 CCTGATGAAGCACCGCGCGCTGACCAAGAGTCAAGAGATGACCTGAGCGCGCTCGCGC 2739

✓ Tue May 23 13:25:57 2006

us-09-555-342b-1.rtf

Page 7

Db 760 CGCCACACACAGCCCAAGGCTGAGCCAGCGCGGCGCTTGAGGCCAGCGACCGCCGGGCGC 703

Search completed: May 22, 2006, 14:56:30
Job ctime : 713.741 secs

This Page Blank (usrp)

|||||
Db 229 CCGGGAAGCTGCTGATGATCAGTTTCAACCACTCAACTCGTGAAGGTGACTAT 288
QY 241 TTTGGCTGAGTTTCCCTATCAAAAAGATCAAGGTGCTGATCTCTTAAACC 300
Db 289 TTTGGCTGAGTTTCCCTATCAAAAAGATCAAGGTGCTGATCTCTTAAACC 348
QY 301 ATTGTGAACAGATTGAAGCCAAAGCAAGTTGTTGAAGTTTGTGTGAATTCTT 360
Db 349 ATTGTGAACAGATTGAAGCCAAAGCAAGTTGTTGAAGTTTGTGTGAATTCTT 408
QY 361 CCGCTGACCAACACAACTCCAAGAACTTCAAGGTACTGTTGCGCTGCAGGTG 420
Db 409 CCGCTGACCAACAACTCCAAGAACTTCAAGGTACTGTTGCGCTGCAGGTG 468
QY 421 AAGCAGATTGGCTCAAGCAGTTGAGGTGAATGACCAAGCCACTCTCTTGAT 480
Db 469 AAGCAGATTGGCTCAAGCAGTTGAGGTGAATGACCAAGCCACTCTCTTGAT 528
QY 481 TCACACATTGCTCAATCTGAGATTGGGGAATTTTGAAGACCTTGAAGACACTTA 540
Db 529 TCACACATTGCTCAATCTGAGATTGGGGAATTTTGAAGACCTTGAAGACACTTA 588
QY 541 GCAAAAATTAATACATACCTCAGCAAGACGCACTAGAGACAAAATCGTGAATTTC 600
Db 589 GCAAAAATTAATACATACCTCAGCAAGACGCACTAGAGACAAAATCGTGAATTTC 648
QY 601 CATAACCAATTGGAACAAACACAGAGATGAGATTTCAAGTCTCTAAGATTGCCGT 660
Db 649 CATAACCAATTGGAACAAACACAGAGATGAGATTTCAAGTCTCTAAGATTGCCGT 708
QY 661 CCGCTAGAGATGATGGAATCCGTTGCAACCGGCCAAGACAGGAGGCAAGATC 720
Db 709 CCGCTAGAGATGATGGAATCCGTTGCAACCGGCCAAGACAGGAGGCAAGATC 768
QY 721 AATCTGGCCGTTGCCAACAGGGAATTCATGTTTCAAGGTTTCAATGAATGATGC 780
Db 769 AATCTGGCCGTTGCCAACAGGGAATTCATGTTTCAAGGTTTCAATGAATGATGC 828
QY 781 TTCAACTGGGCCAAGGTGGGAAAGCTGAGCTTCAAGAGAGGCGTTTCAATCAAGCTC 840
Db 829 TTCAACTGGGCCAAGGTGGGAAAGCTGAGCTTCAAGAGAGGCGTTTCAATCAAGCTC 888
QY 841 CCGCCAGATGCCAATAGTGCCTGACAGAGATACCTTGAATTCCTGATGCGAGT 900
Db 889 CCGCCAGATGCCAATAGTGCCTGACAGAGATACCTTGAATTCCTGATGCGAGT 948
QY 901 TTCTGCAAGTCTTCTGAAAAATCTGTGTTGAACATCATGCTCTTTAGACTTTTGA 960
Db 949 TTCTGCAAGTCTTCTGAAAAATCTGTGTTGAACATCATGCTCTTTAGACTTTTGA 1008
QY 961 GAGCCCAAAACAAAGCCCAAGCCCGTCTTTAGCCGGGGGTCAATTTGGGTTCACT 1020
Db 1009 GAGCCCAAAACAAAGCCCAAGCCCGTCTTTAGCCGGGGGTCAATTTGGGTTCACT 1068
QY 1021 GGTTCGCACTCAGAAAGGTTCTGACTATGTTAAAGAGAGACATAAGAGGTGCA 1080
Db 1069 GGTTCGCACTCAGAAAGGTTCTGACTATGTTAAAGAGAGACATAAGAGGTGCA 1128
QY 1081 TTTGAAAGAGACACAGCAAGATTCATTATCCGAGCCTTGTCAAGCTTACAGA 1140
Db 1129 TTTGAAAGAGACACAGCAAGATTCATTATCCGAGCCTTGTCAAGCTTACAGA 1188
QY 1141 CTGAATTCGGAAGTGTGAGAGCTCAGACAGACAGCCTTACATTTGAGAGAGGT 1200
Db 1189 CTGAATTCGGAAGTGTGAGAGCTCAGACAGACAGCCTTACATTTGAGAGAGGT 1248
QY 1201 GCCGATCTCAGAGGGGCGAGAGCTGCGCGAGAGAGAAAGAAACGAGAGTTTCCGCG 1260
Db 1249 GCCGATCTCAGAGGGGCGAGAGCTGCGCGAGAGAGAAAGAAACGAGAGTTTCCGCG 1308
QY 1261 GAGCCGGGGTCCGACCCGAGCCCTGCGAGAGAGAAAGCCCGGGGTAAACAAGCGCG 1320
|||||

Db 1309 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGTAAACAAGCGCG 1368
QY 1321 GACGAGCGGCTCTGGCGCCCAAGAGAGAAAGAGAGGTGTTAAGATAGAGCCAG 1380
Db 1369 GACGAGCGGCTCTGGCGCCCAAGAGAGAAAGAGAGGTGTTAAGATAGAGCCAG 1428
QY 1381 CAGAGTAAACCTCAGCCCCCGGAGCCAAAGCAAGGCTCTTGATCTGAGCTTCACTT 1440
Db 1429 CAGAGTAAACCTCAGCCCCCGGAGCCAAAGCAAGGCTCTTGATCTGAGCTTCACTT 1488
QY 1441 TCCGAGCTGTCTGTGAATCTCGAGGGGGGAGTGTGCTTGCCTCAAGTACCTTCTCC 1500
Db 1489 TCCGAGCTGTCTGTGAATCTCGAGGGGGGAGTGTGCTTGCCTCAAGTACCTTCTCC 1548
QY 1501 AACCTGAGCCCGACACCAAGAGGCTCTCCCTTGATCAGCCGCTGTAATGACAG 1560
Db 1549 AACCTGAGCCCGACACCAAGAGGCTCTCCCTTGATCAGCCGCTGTAATGACAG 1608
QY 1561 GCTTCCCGGACGAGATGAGAGTGAAGGCGGAGAGAAAGATTCCCACTGATAA 1620
Db 1609 GCTTCCCGGACGAGATGAGAGTGAAGGCGGAGAGAAAGATTCCCACTGATAA 1668
QY 1621 GGTACTTCTATGCTAAGAAATGTTCTACACCGAGCGCACTATCTGAAGATCTCGAA 1680
Db 1669 GGTACTTCTATGCTAAGAAATGTTCTACACCGAGCGCACTATCTGAAGATCTCGAA 1728
QY 1681 GTTATCACTTGTGTGTTCAAGACACAGTGAAGAAAGAGAGCGCATGCCGAGACACTG 1740
Db 1729 GTTATCACTTGTGTGTTTCAAGACACAGTGAAGAAAGAGAGCGCATGCCGAGACACTG 1788
QY 1741 AAAAGTCTATATTCGGAATTTTGAACCTTTGACAAATTTCACTAATTTTCTCAAG 1800
Db 1789 AAAAGTCTATATTCGGAATTTTGAACCTTTGACAAATTTCACTAATTTTCTCAAG 1848
QY 1801 GAAATTTGAGCAACGACTTGGCCCTGTGGGAAGGCGCTCAAAAGCCCAATCAAGATTAC 1860
Db 1849 GAAATTTGAGCAACGACTTGGCCCTGTGGGAAGGCGCTCAAAAGCCCAATCAAGATTAC 1908
QY 1861 CAAGAATCGGCGATGATGATGCTGAAGAACTTCAAGGCAATGAACACACTGCGCTCAC 1920
Db 1909 CAAGAATCGGCGATGATGATGCTGAAGAACTTCAAGGCAATGAACACACTGCGCTCAC 1968
QY 1921 CTGTGAAAGCACAGGAGGCTTGGAGGCGCTTGAAGAAATGAATCAAGGCTCCGCGCG 1980
Db 1969 CTGTGAAAGCACAGGAGGCTTGGAGGCGCTTGAAGAAATGAATCAAGGCTCCGCGCG 2028
QY 1981 CTGAGAACTTCTGAGAGACTTTGAGCTGAGAGAGTGTGTTACTTACCTGCTCAACAC 2040
Db 2029 CTGAGAACTTCTGAGAGACTTTGAGCTGAGAGAGTGTGTTACTTACCTGCTCAACAC 2088
QY 2041 TTCTCTGCGGCGCACTGACACGCGCTCATGCACTAACAAGAGTCTGAGGCGGCTGTGC 2100
Db 2089 TTCTCTGCGGCGCACTGACACGCGCTCATGCACTAACAAGAGTCTGAGGCGGCTGTGC 2148
QY 2101 AAACACACCGCGCGAGCCACGCGACTTCAAGGACTGCGCGAGCCGTTTGGCAAGATC 2160
Db 2149 AAACACACCGCGCGAGCCACGCGACTTCAAGGACTGCGCGAGCCGTTTGGCAAGATC 2208
QY 2161 ACGAGAGTGTGGCAACACTTCAAGGTACAGATGAACAAGTGAAGATTTCCAGAAGCTG 2220
Db 2209 ACGAGAGTGTGGCAACACTTCAAGGTACAGATGAACAAGTGAAGATTTCCAGAAGCTG 2268
QY 2221 CACGAACCTCAAGAAAGATTGATTTGGCAATTGCAATCTTGTGTTTCCGGAAAGGAGTTC 2280
Db 2269 CACGAACCTCAAGAAAGATTGATTTGGCAATTGATTTGCAATCTTGTGTTTCCGGAAAGGAGTTC 2328
QY 2281 ATCGTCTGGGAGGCTTCAAGAGCTCTCGGGAAAGGGCTTCAAGACGCAATGTTCTTC 2340
Db 2329 ATCGTCTGGGAGGCTTCAAGAGCTCTCGGGAAAGGGCTTCAAGACGCAATGTTCTTC 2388
QY 2341 CTGTTCAACAGCTCTGCTATACAGAGCGGGGGCTGACGCGCTTCAATGATTTTAA 2400
Db 2389 CTGTTCAACAGCTCTGCTATACAGAGCGGGGGCTGACGCGCTTCAATGATTTTAA 2448
|||||

QY 2401 GTCCAGGAGGAGTCCCGCTCTATGGCATGACGATTGAGAGAGCGAAGACGAGTGGGG 2460
DB 2449 GTCCAGGAGGAGTCCCGCTCTATGGCATGACGATTGAGAGAGCGAAGACGAGTGGGG 2508
QY 2461 GTGCCCCACTGCTGACCTCCGCGGCGACGCGAGTGCATCATCTGCGCGCAGTCTT 2520
DB 2509 GTGCCCCACTGCTGACCTCCGCGGCGACGCGAGTGCATCATCTGCGCGCAGTCTT 2568
QY 2521 CGGTCCGAGATGAGAGAGTGGGTGAGGACATCCAGATGGCCATTGACTTGGCGAGAG 2580
DB 2569 CGGTCCGAGATGAGAGAGTGGGTGAGGACATCCAGATGGCCATTGACTTGGCGAGAG 2628
QY 2581 AGCAGAGAGCCCGCCCTGAGTTCCTGCGCAGAGAGCCCGCTGACAAAGTCCCTGAT 2640
DB 2629 AGCAGAGAGCCCGCCCTGAGTTCCTGCGCAGAGAGCCCGCTGACAAAGTCCCTGAT 2688
QY 2641 GAAGCCACGCGCGCTGACAGAGATGACCTGAGCGCTTCGCGCAGATCGCTG 2700
DB 2689 GAAGCCACGCGCGCTGACAGAGATGACCTGAGCGCTTCGCGCAGATCGCTG 2748
QY 2701 GAGCGCAGGCGCCCGCAGCCGCGCAACATGCTGCTGCGCAGCGCAACACC 2760
DB 2749 GAGCGCAGGCGCCCGCAGCCGCGCAACATGCTGCTGCGCAGCGCAACACC 2808
QY 2761 AGCGTCCATGGTGAAGTTCAGATCGAGTGAAGATGAGTTCGAGAACTGCTG 2820
DB 2809 AGCGTCCATGGTGAAGTTCAGATCGAGTGAAGATGAGTTCGAGAACTGCTG 2868
QY 2821 AGGAAATTCAAAAACAGCAACGCGGTGAGAGAGTGTGGTGTTCACAACTTCTG 2880
DB 2869 AGGAAATTCAAAAACAGCAACGCGGTGAGAGAGTGTGGTGTTCACAACTTCTG 2928
QY 2881 CTGTTTCTTACAAATACACCGAGACATCATCCCTTCCAGCTGCTTCTGCTGGC 2940
DB 2929 CTGTTTCTTACAAATACACCGAGACATCATCCCTTCCAGCTGCTTCTGCTGGC 2988
QY 2941 TACTGCTACCACTCCCTCTGAGTCCGAGACATCCAGAAAGACTTACAGTTCAGAGCTG 3000
DB 2989 TACTGCTACCACTCCCTCTGAGTCCGAGACATCCAGAAAGACTTACAGTTCAGAGCTG 3048
QY 3001 CACTTCAAGTCCGACGCTTACTTACAGGCGGAGAAAGCAGATACAGTTCGAAAGGTG 3060
DB 3049 CACTTCAAGTCCGACGCTTACTTACAGGCGGAGAAAGCAGATACAGTTCGAAAGGTG 3108
QY 3061 ATGGAGATGATCCGACAGTGCACCAAGCTTGCCTCGGAGCCCACTGTTGAGCCACAA 3120
DB 3109 ATGGAGATGATCCGACAGTGCACCAAGCTTGCCTCGGAGCCCACTGTTGAGCCACAA 3168
QY 3121 GAGTCTCTTGTAT 3135
DB 3169 GAGTCTCTTGTAT 3183

RESULT 2
AB008430 3442 bp mRNA linear PRI 13-FEB-1999
LOCUS AB008430 Homo sapiens mRNA for CDEP, complete cds.
DEFINITION
AB008430
ACCESSION
AB008430.1 GI:2766164
VERSION
CDEP.
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE
AUTHORS
TITLE
Koyano, Y., Kawamoto, T., Shen, M., Yan, W., Noshiro, M., Fujii, K. and
Kato, Y.
Molecular cloning and characterization of CDEP, a novel human
protein containing the ezrin-like domain of the band 4.1
superfamily and the Dbl homology domain of Rho guanine nucleotide
exchange factors

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
Biochem. Biophys. Res. Commun. 241 (2), 369-375 (1997)
9425278
2 (base 1 to 3442)
Koyano, Y., Kawamoto, T., and Kato, Y.
Direct Submission
Submitted (22-OCT-1997) Takeshi Kawamoto, Hiroshima University
School of Dentistry, Department of Biochemistry, 1-2-3 Kasumi
Minami-Ku, Hiroshima, Hiroshima 734, Japan
(E-mail: tkawamo@ipc.hiroshima-u.ac.jp, Tel: 082-257-5688,
Fax: 082-257-5629)

FEATURES
source
Location/Qualifiers
1..3442
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="chondrocyte"
/tissue_type="cartilage"
/dev_stage="embryo"
49..3186
/function="Rho Guanine Nucleotide Exchange Factor"
/note="Band 4.1 superfamily"
/codon_start=1
/product="CDEP"
/protein_id="BAA24267.1"
/db_xref="GI:2766165"

CDS
/translation="MGETIRORPTPGSRILGAPRNSGISTLERGQKPPPTSGKLVSTIKI
QMLDITQENAFVPGARPGVLDVAVCNHNLVBYGDFGLFEPDHKITTWLDLKRIV
KQIRKPHVAVFVAFPPDHTQLDSELRFLFALQVQDLAQGLTCNDLSALLI
SHIVSEIDFDBALDREHLANKYIPQDADLEDKIVERHNNHIGOTPAESPQLEI
ARLEWYGRILHPADOREGTXINLAVANTGIVFOGFTINAFNAVKRLSEKRF
LILKRPDASAVODTLEFLMASRDFCKSPWKI CVEHAFPLPEEKPKPKVLPFRG
SSPRPSGRQKOVLDVKEGHHKQVPEEKHKTSHIRSLAQPELNEBEVLEOSOS
TSITTBGABESRGSGCRGKPKYSAGSPGSHPSAPRPSRPNKQDAGASAPTEE
EEBVADKRTQSKPQPPQSTGLTSPHLSVSVQGVAPANTVLSPNLSDTKQ
ASPLISPLINDQACPTDDEGRRRRFPDIAVFAKSVSTTERVYLDLEVTISWF
OSTVSKEDPALASLIPNPEPLHKPTNFAKEIQEALAMEGSSNOIDRYORIG
DVLKNIQGMKLAHLIMGSEALELNGIKSSRLNENFCDFELQKCYLPANFL
LPIRLHMYKQVLEBELCKHHPNSHADPDCRAALAEITEMVAOLHGMVKNENPQKL
HEIKDLIGIDNLVYDPREFTIRLGSLSKLSGGLQORMFPLPNDVLLYTSRGTASNO
FVHGQPLVYGNITIESEDEMGVPHCLTLRGQOSTITVAASRSEMEKVEDIQMAD
LAEKSSVPAEFLASPPDNKSPPDEXTAADQSEBDLSRSTSLERQAFHRGNATVAV
CHNRNYSVMPVDFIASVENOLSGNLRRKNSNGQKLVAVVTNFCLPYKSHQDHP
LASLPLGLSLTIPSEBNIQDYVFKLHFKSHVYFFRAESEYTERMEVIRSATSS
ASRPVPLSHKESLVY"

polyA_site
3442
/note="50 a nucleotides"

ORIGIN
Query Match 100.0%; Score 3135; DB 5; Length 3442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAGAAATAGACAGAGCCGACCCGACGATCAGACTGGGGCCCCGAGAAATTGC 60
DB 49 ATGGGAGAAATAGACAGAGCCGACCCGACGATCAGACTGGGGCCCCGAGAAATTGC 108
QY 61 GGGATCAGTACTCTTGGAAAGTGGACGAAAGCCGCCCAACACCTTCAGAGAAACTCGTG 120
DB 109 GGGATCAGTACTCTTGGAAAGTGGACGAAAGCCGCCCAACACCTTCAGAGAAACTCGTG 168
QY 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAGTTCACAAAGAGCT 180
DB 169 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAGTTCACAAAGAGCT 228
QY 181 CTTGGGAGAGTGTCTGTGATGACAGTTTCAACCACTCAACTCTGTGAGAGTGAATAT 240
DB 229 CTTGGGAGAGTGTCTGTGATGACAGTTTCAACCACTCAACTCTGTGAGAGTGAATAT 288
QY 241 TTTGGCTCGAGTTTCTTGTATCAGAAAGATCAGAGTGTGCTGATCTCTTAAACCC 300
DB 289 TTTGGCTCGAGTTTCTTGTATCAGAAAGATCAGAGTGTGCTGATCTCTTAAACCC 348
QY 301 ATTGTAAACAGATTGAGAGCCAAAGCAGTGTGTTGAAGTTGTGTGAAATTCTTT 360

Db 349 ATGTGAAACAGATTGGAAGCCAAAGCAGTTGTGTTAAAGTTGTGTGAATCTTT 408
Qy 361 CCGCTGACCAACACAACTCCAGAGAAATCAACAAGTACCTGTTCGCGCTGACAGTG 420
Db 409 CCGCTGACCAACAACTCCAGAGAAATCAACAAGTACCTGTTCGCGCTGACAGTG 468
Qy 421 AAGCAGACTTGGCTCAAGGACAGTGTATATGACACAGCAGCTCTCTTGAT 480
Db 469 AAGCAGACTTGGCTCAAGGACAGTGTATATGACACAGCAGCTCTCTTGAT 528
Qy 481 TCACACATTTGTGCAATCTGAGATTGGGGATTTTGAAGACCTTGGACAGAGCACTTA 540
Db 529 TCACACATTTGTGCAATCTGAGATTGGGGATTTTGAAGACCTTGGACAGAGCACTTA 588
Qy 541 GGAATAAATAATATACCTACCTGACAGAAAGCAGCTAGAGAGCAAAATCGTGAATTTTAC 600
Db 589 GGAATAAATAATATACCTACCTGACAGAAAGCAGCTAGAGAGCAAAATCGTGAATTTTAC 648
Qy 601 CATTAACCACTTGGACAAACACAGAGATCAGATTTTCCAGCTCCTAGAGATTGCCGT 660
Db 649 CATTAACCACTTGGACAAACACAGAGATCAGATTTTCCAGCTCCTAGAGATTGCCGT 708
Qy 661 CCGCTAGAGATGTATGGAATCCGGTTGCAACCGGCGCAAGGACAGGGAGCAGAAAGATC 720
Db 709 CCGCTAGAGATGTATGGAATCCGGTTGCAACCGGCGCAAGGACAGGGAGCAGAAAGATC 768
Qy 721 AATCTGCGCGTTGCCCAACCGGAAATCTTAAGTTTCAAGGTTTCAATTAAGTCAATGCC 780
Db 769 AATCTGCGCGTTGCCCAACCGGAAATCTTAAGTTTCAAGGTTTCAATTAAGTCAATGCC 828
Qy 781 TTCAACTGGGCGCAAGGTGCGGAAAGTGAAGTCTTAAGAGAAAGCGCTTCTCATCAAGCTC 840
Db 829 TTCAACTGGGCGCAAGGTGCGGAAAGTGAAGTCTTAAGAGAAAGCGCTTCTCATCAAGCTC 888
Qy 841 CCGCCAGATGCCAATATGTCGTAACAGATACCTTGAATTCCTGATGCCAGTCGGAGT 900
Db 889 CCGCCAGATGCCAATATGTCGTAACAGATACCTTGAATTCCTGATGCCAGTCGGAGT 948
Qy 901 TTCTGCAATCTCTTGGAAATCTGTGTGAAATCAATCAAGCCTCTTGAAGCTTTTGA 960
Db 949 TTCTGCAATCTCTTGGAAATCTGTGTGAAATCAATCAAGCCTCTTGAAGCTTTTGA 1008
Qy 961 GAGCCCAACCAAAAGCCCAAGCCGTCCTTTAGCCGGGGCATCATTTGCGTTCAGT 1020
Db 1009 GAGCCCAACCAAAAGCCCAAGCCGTCCTTTAGCCGGGGCATCATTTGCGTTCAGT 1068
Qy 1021 GGTCCGACTCAAGAGAGGTTCTCGACTATGTTAAAGAGAGACATAAGAGGTGCAG 1080
Db 1069 GGTCCGACTCAAGAGAGGTTCTCGACTATGTTAAAGAGAGACATAAGAGGTGCAG 1128
Qy 1081 TTTTGAAGAGACAGCAAGATTCATTTATCCGAGCCTTGCTTCAAGCTACAGAA 1140
Db 1129 TTTTGAAGAGACAGCAAGATTCATTTATCCGAGCCTTGCTTCAAGCTACAGAA 1188
Qy 1141 CTGAATTCGAAAGTCTGAGAGAGTCTCAGCAGAGCAGCCTTACATTTGGAGAGGT 1200
Db 1189 CTGAATTCGAAAGTCTGAGAGAGTCTCAGCAGAGCAGCCTTACATTTGGAGAGGT 1248
Qy 1201 GGTCAATCTTCAAGGGGGCAGAGCTGCGGCGAGAGAAAGAACCGAAGGTTTCCGCGGG 1260
Db 1249 GGTCAATCTTCAAGGGGGCAGAGCTGCGGCGAGAGAAAGAACCGAAGGTTTCCGCGGG 1308
Qy 1261 GAGCCGGGGTCCAGCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTAAACAAGCAGGGG 1320
Db 1309 GAGCCGGGGTCCAGCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTAAACAAGCAGGGG 1368
Qy 1321 GACGAGACCGCTCGCGGCCCAAGAGAGAGAGAGAGAGTCTTAAAGATAGAACCCAG 1380
Db 1369 GACGAGACCGCTCGCGGCCCAAGAGAGAGAGAGAGAGTCTTAAAGATAGAACCCAG 1428
Qy 1381 CAGAGTAAACTTCAGCTCCCGCAGCCAGACAGGCTCCCTGACTGGCAGTCTCACTTT 1440

Db 1429 CAGAGTAAACTTCAGCCCCCGCAGCCAAAGCAGAGCTCCCTGACTGGCAGTCTCACTT 1488
Qy 1441 TCCGAGCTGTCTGTAACTCGCAGGGGGAGAGTGGCCCTGCCCAAGCTTGTCTCC 1500
Db 1489 TCCGAGCTGTCTGTAACTCGCAGGGGGAGAGTGGCCCTGCCCAAGCTTGTCTCC 1548
Qy 1501 AACCTGAGCCCCCAGCAGCAAGAGGCTCTCCCTTGAATCAGCCGCTGTGAATGACAG 1560
Db 1549 AACCTGAGCCCCCAGCAGCAAGAGGCTCTCCCTTGAATCAGCCGCTGTGAATGACAG 1608
Qy 1561 GCTTCCCCCGCAGCAGAGATGAGAGCCGAGAGAAAGATTCCTCACTGATTA 1620
Db 1609 GCTTCCCCCGCAGCAGAGATGAGAGCCGAGAGAAAGATTCCTCACTGATTA 1668
Qy 1621 GGTACTTCAATAGCTAAGAAAGTCTTACACAGAGCAATATCTGAGAGTCTGAA 1680
Db 1669 GGTACTTCAATAGCTAAGAAAGTCTTACACAGAGCAATATCTGAGAGTCTGAA 1728
Qy 1681 GTTATCACTTGTGGTTTCAAGACAGTGAAGAAAGAGAGCCATGCGGAAGCATG 1740
Db 1729 GTTATCACTTGTGGTTTCAAGACAGTGAAGAAAGAGAGCCATGCGGAAGCATG 1788
Qy 1741 AAAAGTCTATTTCCCGAATTTTGAACCTTTGCAAAATTTTCAATTAATTTTCTCAAG 1800
Db 1789 AAAAGTCTATTTCCCGAATTTTGAACCTTTGCAAAATTTTCAATTAATTTTCTCAAG 1848
Qy 1801 GAAATTTGAGCAAGACTTGCCCTGTGGGAAGGCGCTCAATAGCCCAATCAGAGTTAC 1860
Db 1849 GAAATTTGAGCAAGACTTGCCCTGTGGGAAGGCGCTCAATAGCCCAATCAGAGTTAC 1908
Qy 1861 CAAAGATTCGGGATGTCAATGCTGAAGAACATTCAGGGCATGAAGCACTGGCGCTCAC 1920
Db 1909 CAAAGATTCGGGATGTCAATGCTGAAGAACATTCAGGGCATGAAGCACTGGCGCTCAC 1968
Qy 1921 CTGTGGAAGCACAGGAGGCTTGAAGGCCCTGAGAGAAATGAAATCAAGGCTCCGCGCG 1980
Db 1969 CTGTGGAAGCACAGGAGGCTTGAAGGCCCTGAGAGAAATGAAATCAAGGCTCCGCGCG 2028
Qy 1981 CTGAGAACTTCTGAGAGACTTTGAGCTGCAGAGAGTGTGTTACTACCGCTCAACACC 2040
Db 2029 CTGAGAACTTCTGAGAGACTTTGAGCTGCAGAGAGTGTGTTACTACCGCTCAACACC 2088
Qy 2041 TTCTCTTGGCGCACTGACCGGCTCATGCAACAAGAGTCTTGAAGCGGCTGTGC 2100
Db 2089 TTCTCTTGGCGCACTGACCGGCTCATGCAACAAGAGTCTTGAAGCGGCTGTGC 2148
Qy 2101 AAAACACACCCCGCAGAGCAGCGCACTTCAAGGAGCTGCGGACGCTTTGGCAGAGATC 2160
Db 2149 AAAACACACCCCGCAGAGCAGCGCACTTCAAGGAGCTGCGGACGCTTTGGCAGAGATC 2208
Qy 2161 ACGAGATGAGTGGCAGAGCTCCAGCTCAGATGATCAAGATGAGAAATTTCCAGAGCTG 2220
Db 2209 ACGAGATGAGTGGCAGAGCTCCAGCTCAGATGATCAAGATGAGAAATTTCCAGAGCTG 2268
Qy 2221 CACGAATCTCAAGAAAGATTTGATGGAATCTTGTGTTCCGGGAAGGAGTTT 2280
Db 2269 CACGAATCTCAAGAAAGATTTGATGGAATCTTGTGTTCCGGGAAGGAGTTT 2328
Qy 2281 ATCCGTCTGGGAGAGCTCAGCAAGCTTCGCGGAAAGGGCTCACAAGCGCAATGTTCTTC 2340
Db 2329 ATCCGTCTGGGAGAGCTCAGCAAGCTTCGCGGAAAGGGCTCACAAGCGCAATGTTCTTC 2388
Qy 2341 CTGTTCAAGAGAGTCTGCTATACAGAGCCGGGGCTGACGGCTCCATCAGTTTAA 2400
Db 2389 CTGTTCAAGAGAGTCTGCTATACAGAGCCGGGGCTGACGGCTCCATCAGTTTAA 2448
Qy 2401 GTTCAAGGAGAGTCTCCGCTTATGAGCATGATTTGAGAGAGCGAAGACAGATGAGGGG 2460
Db 2449 GTTCAAGGAGAGTCTCCGCTTATGAGCATGATTTGAGAGAGCGAAGACAGATGAGGGG 2508
Qy 2461 GTTCCCACTGCTGACCTCTCGGGGCGAGCGGCGAGTCCATATGTTGCGCCGAGTTT 2520
Db 2509 GTTCCCACTGCTGACCTCTCGGGGCGAGCGGCGAGTCCATATGTTGCGCCGAGTTT 2568

2521 CGGTCCGAGATGAGAGAGTGGGTTGAGGACATCCAGATGCGCATTCAGCTGGCGGAGAG 2580
2569 CGGTCCGAGATGAGAGAGTGGGTTGAGGACATCCAGATGCGCATTCAGCTGGCGGAGAG 2628
2581 AGCAGCAGCCCGCCCTGAGTTCCTGGCAGCAGCCCGCCCTGAGCAACAAGTCCCTGAT 2640
2629 AGCAGCAGCCCGCCCTGAGTTCCTGGCAGCAGCCCGCCCTGAGCAACAAGTCCCTGAT 2688
2641 GAAGCCACCGCGGCTGACAGAGTCAAGAGATGACCTGAGCCCTTCGCGCAGATGCTG 2700
2689 GAAGCCACCGCGGCTGACAGAGTCAAGAGATGACCTGAGCCCTTCGCGCAGATGCTG 2748
2701 GAGCGCAGGCGCCCGCAGCAGCGGCAACAATGAGTGCAGTGGGCGCAGCGCAACACC 2760
2749 GAGCGCAGGCGCCCGCAGCAGCGGCAACAATGAGTGCAGTGGGCGCAGCGCAACACC 2808
2761 AGCGTCTCATGAGTGAATTCAGATCGCAGTGGAGATGATGTTGTTGAGAACTGCTG 2820
2809 AGCGTCTCATGAGTGAATTCAGATCGCAGTGGAGATGATGTTGTTGAGAACTGCTG 2868
2821 AGGAAATTCAAAAACAGCAACGGGTGGCAGAGCTGTGGTGTGTTACAAACTTCTGC 2880
2869 AGGAAATTCAAAAACAGCAACGGGTGGCAGAGCTGTGGTGTGTTACAAACTTCTGC 2928
2881 CTGTTCTTCTACAAATCACACACGAGCAATCATCCCTTGGCCAGCCCTGCTGCTCGC 2940
2929 CTGTTCTTCTACAAATCACACGAGCAATCATCCCTTGGCCAGCCCTGCTGCTGCTGCGC 2988
2941 TACTGCTCAGCATCCCTCTGTAGTCCGAGAAATCCAGAAATGATGTTCAAGCTG 3000
2989 TACTGCTCAGCATCCCTCTGTAGTCCGAGAAATCCAGAAATGATGTTCAAGCTG 3048
3001 CACTTCAAGTCCAGCTTACTTACTTACGGGCGGAAAGCGAGTACAGTTGGAAGGTG 3060
3049 CACTTCAAGTCCAGCTTACTTACTTACGGGCGGAAAGCGAGTACAGTTGGAAGGTG 3108
3061 ATGGAAGTATCCGCGAGTGCACACAGCTCTGCTCCGACCCGACGTTGAGACCAAA 3120
3109 ATGGAAGTATCCGCGAGTGCACACAGCTCTGCTCCGACCCGACGTTGAGACCAAA 3168
3121 GAGTCTCTGTGTAT 3135
3169 GAGTCTCTGTGTAT 3183

RESULT 3
BC041595 3546 bp mRNA 1linear PRI 08-MAR-2005
LOCUS
DEFINITION Homo sapiens FERM, RhoGEF (ARHGEF) and plectectrin domain protein 1
(Chondrocyte-derived), transcript variant 1, mRNA (cDNA clone
MGC:52223 IMAGE:5443230), complete cds.
ACCESSION BC041595
VERSION BC041595.1 GI:27370806
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 3546)
Struhsberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
Collins FS, Wagner L, Shenmen CM, Schuler GD, Altshul SF, Zeeberg
B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin
GM, Hong L, Stapleton M, Soares MB, Bonaldo MP, Casavant TL,
Schetz TE, Brownstein WJ, Ueda TB, Toshiyuki S, Carninci P,
Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullaby
SJ, Bosak SA, McGowan PJ, McKernan KJ, Malek JA, Gunaratne PH,
Richard S, Worley KC, Hale S, Garcia AM, Gay LJ, Holik SM,
Villalón DK, Wuzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahy J,
Helton E, Kesteman M, Madan A, Rodrigues S, Sanchez A, Whiting M,
Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakeley RW,

Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J,
Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skelton U,
Smalhus DE, Schmechel A, Schein JE, Jones SJ and Marra MA.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3546)
Director MGC Project.
Direct Submission
Submitted (20-DEC-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayle, K., Beckstrom-Stenberg, S.M., Benjamin, B.,
Blakeley, R.W., Bouffard, G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Larit, P., Legaspi, R.,
Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McLokey, J.C.,
McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W.,
Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRRL Plate: 44 Row: 1 Column: 24.
Location/Qualifiers
1. 3546
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:52223 IMAGE:5443230"
/issue_type="Brain, astrocytoma, grade IV"
/clone_1b="NIH_MGC_98"
/lab_host="DH10B"
/note="Vector: pOTB7"
1. 3546
/gene="PARP1"
/note="Synonym: CDEP, MGC87400"
/db_xref="GeneID:10160"
/db_xref="MIM:602654"
266..3403
/gene="PARP1"
/codon_start=1
/product="FERM, RhoGEF, and plectectrin domain protein 1,
isoform 1"
/protein_id="AAH41595.1"
/db_xref="GeneID:10160"
/db_xref="MIM:602654"
/translation="MEIRQRTPGSRGLGAPNSGISTIERGOKPEPTPSGTLVSTIKI
QMLDTPQAEVQAPGQVLDVAVCNHNLVBDYFGLEFPDHRKITVWLDLKITV
KQIRPKYVAVVAVKPEPDDTOLOEILRYLFLAQVQDLAOGGLTNDTSALILI
SHIYSEIDPEADLREHLAKNKYIPODADLEKIVFHHNHIGTPAESFOLLEI
ARLEMYGRIHLPADREKTKNLAVANTGILVFOGFTKINAFNAKVKLSFKRKF
LKLKPDANSAYOPTLEPMASRDCKSPFKICVHHAFPLFEEBKPKPKVLSRGS
SFRSGRQKOVLDYKVGCGKTKQAPRKHSHKISIRLSAOPTEINSEVLEQSOO
TSLTREGBASFPQSGCRGKRPKYSAGPSGSPAPRPSRPAKNGKADGAAAPTEE
BEKVVDKQOSKPPQPPSTSLTSGPLSLSVNSQGVAPAVNTLSPLSPDTS
ASPLSLPLNDACPTDDEGRRRKPPTDRAVFAKESVSTTERLYLNDLIVTISWF
GSTVSKEDAMPALKSLIPNPFPLHKFTFNLKEIEQLALMEGRSNAQIRDYQIG

DVLMKNIQGMKHLAAHLWKHSBALBLENKISRSRLNENFCRDFELOKVCYLPNTFL
LAEHLRMHYKQVLERLCGHPSPSHADPFDCAPALAEITBEMVAQHGTMKENTOKL
HELKDLGIDNLVYGRREFIRLSXIGLOMRFEINDVLATYSRRLTASNO
FVYHQDLPYGMGTIESEBDEKVPKICLTIRGOROSIYVAASRSMEKREVEDIQMAID
LAEKSSPAEPFLASPDPNKSPEDEATADDESBDLSRSRTSLRQAPHRANTVHV
CWHRTSVSMVDFSIJAVENQLSRKRNKNGWQKLVVNTNLCFLPYKSHQONHP
LASLPDLYSLTIPSEBENIQRDVFYKLFKSHVYVFRASEYTEREWMVEVRSATSS
ASRPHVLSHKEISLVY"

ORIGIN

Query Match 100.0%; Score 3135; DB 5; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGAGAAATAGAGAGAGCGCCAGATCCAGATCAGACTGGGGGCCCCGGAAATTGG 60
DB 266 AATGGGAAATAGAGAGAGCGCCAGATCCAGATCAGACTGGGGGCCCCGGAAATTGG 325
QY 61 GGGATCAGTACCTTGGGAGCTGGACAGAAAGCCGCCCAACACTTCAGAGAAACCTCGTG 120
DB 326 GGGATCAGTACCTTGGGAGCTGGACAGAAAGCCGCCCAACACTTCAGAGAAACCTCGTG 385
QY 121 TCATATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAGTTCCAAAGAGCT 180
DB 386 TCATATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAGTTCCAAAGAGCT 445
QY 181 CCGGGAGAGGTGCTGATGAGTTGCAACCACTCAACCTGCTGGAGGTGACTAT 240
DB 446 CCGGGAGAGGTGCTGATGAGTTGCAACCACTCAACCTGCTGGAGGTGACTAT 505
QY 241 TTTGGCTCGAGTTTCTGATCACAAAAAGATCACCGGTGCTGATCTCTAAAAACC 300
DB 506 TTTGGCTCGAGTTTCTGATCACAAAAAGATCACCGGTGCTGATCTCTAAAAACC 565
QY 301 ATTGGAACAGATTGAAGGCCAAAGCACTGTTGTTTAAGTTGTGGTGAATCTTT 360
DB 566 ATTGGAACAGATTGAAGGCCAAAGCACTGTTGTTTAAGTTGTGGTGAATCTTT 625
QY 361 CGGCGTACACACCACTCCAGAGAAACTCACAAAGGTACTGTTGCGCTGACGAGTG 420
DB 626 CGGCGTACACACCACTCCAGAGAAACTCACAAAGGTACTGTTGCGCTGACGAGTG 685
QY 421 AAGCAGAGCTTGGCTCAAGGAGGTTGACGTGTAATGACACAGCGCAGCTCTTTGATT 480
DB 686 AAGCAGAGCTTGGCTCAAGGAGGTTGACGTGTAATGACACAGCGCAGCTCTTTGATT 745
QY 481 TCACACATTTGCAATCTGAGATTTGGGGATTTTGAATGAAGCAACAGCGCAGCTT 540
DB 746 TCACACATTTGCAATCTGAGATTTGGGGATTTTGAATGAAGCAACAGCGCAGCTT 805
QY 541 GCAAAAATAAATPACTACTCAGCAAGAGCGCACTGAGAGACAAATGGTGGAAATTTCAC 600
DB 806 GCAAAAATAAATPACTACTCAGCAAGAGCGCACTGAGAGACAAATGGTGGAAATTTCAC 865
QY 601 CATTAACACATTTGACAAACACCAAGCAGATCAGATTTTCAGCTCTTAGAGATTGCCGT 660
DB 866 CATTAACACATTTGACAAACACCAAGCAGATCAGATTTTCAGCTCTTAGAGATTGCCGT 925
QY 661 CGGCTAGAGATGTATGGAAATCCGGTTGGCACCGGCGCAAGAGAGGAGGACGAAAGATC 720
DB 926 CGGCTAGAGATGTATGGAAATCCGGTTGGCACCGGCGCAAGAGAGGAGGACGAAAGATC 985
QY 721 AATCTGGCGCTTGGCAACCGGGAATTTAGTGTTCAGGGTTTCTAAGATCAATGCGC 780
DB 986 AATCTGGCGCTTGGCAACCGGGAATTTAGTGTTCAGGGTTTCTAAGATCAATGCGC 1045
QY 781 TTCAACTGGGCCAAGGTGCGAAGCTGAGCTTCAAGAGAAAGCGCTTTCTCACTCAAGCTC 840
DB 1046 TTCAACTGGGCCAAGGTGCGAAGCTGAGCTTCAAGAGAAAGCGCTTTCTCACTCAAGCTC 1105
QY 841 CGGCGAGATGCCAATATGTCGCTACAGAGTACCTTGGAAATTCCTGATGGCCAGTGGGAT 900
DB 1106 CGGCGAGATGCCAATATGTCGCTACAGAGTACCTTGGAAATTCCTGATGGCCAGTGGGAT 1165

QY 901 TTCTCAAGTCTCTTGGAAAAATCTGTGTGAACATCATGCTTCTTTAGACTTTTGA 960
DB 1166 TTCTCAAGTCTCTTGGAAAAATCTGTGTGAACATCATGCTTCTTTAGACTTTTGA 1225
QY 961 GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCCGGGGTCAATCTTTGGTTCACT 1020
DB 1226 GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCCGGGGTCAATCTTTGGTTCACT 1285
QY 1021 GGTGGACTCGAAGCAGAGTTCTGACTATGTTAAAGAGGAGCATTAAGAAAGGTGACG 1080
DB 1286 GGTGGACTCGAAGCAGAGTTCTGACTATGTTAAAGAGGAGCATTAAGAAAGGTGACG 1345
QY 1081 TTTGAAAGGAGCAGACGAAAGTTCTATTCGAGACCTTGTCTTACAGCTTACAGA 1140
DB 1346 TTTGAAAGGAGCAGACGAAAGTTCTATTCGAGACCTTGTCTTACAGCTTACAGA 1405
QY 1141 CTGAATTCGAAAGTCTGAGCAGTCTTCAGACAGACCAAGCTTACATTTGAGAAAGT 1200
DB 1406 CTGAATTCGAAAGTCTGAGCAGTCTTCAGACAGACCAAGCTTACATTTGAGAAAGT 1465
QY 1201 GCCGAATCTCCAGGGGGGCGAGAGTCCGCGAGAGAAAGGAAACGAAAGTTCCCGCGGG 1260
DB 1466 GCCGAATCTCCAGGGGGGCGAGAGTCCGCGAGAGAAAGGAAACGAAAGTTCCCGCGGG 1525
QY 1261 GAGCCGGGGTCCGACCCGAGCCTTGCAGAGAGAAAGCCCGCGGTTAACAGAGGCG 1320
DB 1526 GAGCCGGGGTCCGACCCGAGCCTTGCAGAGAGAAAGCCCGCGGTTAACAGAGGCG 1585
QY 1321 GAGGAGCCGCTCGGCGCCCAAGAGAGAGAGAGAGTCTTTAAGATTGAGACCCAG 1380
DB 1586 GAGGAGCCGCTCGGCGCCCAAGAGAGAGAGAGAGTCTTTAAGATTGAGACCCAG 1645
QY 1381 CAGAGTAAACCTCAGCCCCCGAGCCAGACAGGCTCCTGACCTGGAGTCTCACTT 1440
DB 1646 CAGAGTAAACCTCAGCCCCCGAGCCAGACAGGCTCCTGACCTGGAGTCTCACTT 1705
QY 1441 TCAGAGCTGTCTGTGATCTCGAGGGGGAGTGGCCCTGCAACGTGACTGTCTCCC 1500
DB 1706 TCAGAGCTGTCTGTGATCTCGAGGGGGAGTGGCCCTGCAACGTGACTGTCTCCC 1765
QY 1501 AACCTGAGCCCGGACACAGAGAGGCTCTCTTGAATCAGCCGCTGTGAATGACAG 1560
DB 1766 AACCTGAGCCCGGACACAGAGAGGCTCTCTTGAATCAGCCGCTGTGAATGACAG 1825
QY 1561 GCTGCCCCCGGAGGAGATGAGGAGTGGGGCGGAGAGAGATTCCTCAACTGATAA 1620
DB 1826 GCTGCCCCCGGAGGAGATGAGGAGTGGGGCGGAGAGAGATTCCTCAACTGATAA 1885
QY 1621 GCGTACTTCAATGACTAAGGAAGTGTCTACACCGAGCGAACATATCTGAAGATCTCGAA 1680
DB 1886 GCGTACTTCAATGACTAAGGAAGTGTCTACACCGAGCGAACATATCTGAAGATCTCGAA 1945
QY 1681 GTTATCATCTTGTGTTTCAAGACAGATGAGCAAAAGAGAGCCATGCGGAGACATG 1740
DB 1946 GTTATCATCTTGTGTTTCAAGACAGATGAGCAAAAGAGAGCCATGCGGAGACATG 2005
QY 1741 AAAAGTCTCATATTTCCGAAATTTTGAACCTTTGCAAAATTCATATTAATTTTCTCAAG 1800
DB 2006 AAAAGTCTCATATTTCCGAAATTTTGAACCTTTGCAAAATTCATATTAATTTTCTCAAG 2065
QY 1801 GAAATTGAGCAACGACTTGCCCTGTGGAAAGGCCGCTCAAAATGCCAAATCAGAGATTAC 1860
DB 2066 GAAATTGAGCAACGACTTGCCCTGTGGAAAGGCCGCTCAAAATGCCAAATCAGAGATTAC 2125
QY 1861 CAAGAATCGGAGATGTCACTGCTGAAGAACATTCAGGCGATGAGACACTTGGCGCTCAC 1920
DB 2126 CAAGAATCGGAGATGTCACTGCTGAAGAACATTCAGGCGATGAGACACTTGGCGCTCAC 2185
QY 1921 CTGTGGAAGCAGAGAGGCTTGGAGGCTTGGAGAAATGGAATGAAGAGCTCCGCGCGG 1980
DB 2186 CTGTGGAAGCAGAGAGGCTTGGAGGCTTGGAGAAATGGAATGAAGAGCTCCGCGCGG 2245

1981 CTGAGAACTTCTGACAGAGATTGAGCTGACAGAGGTTGTACTTACCTACCGCTCAACACC 2040
2246 CTGAGAACTTCTGACAGAGATTGAGCTGACAGAGGTTGTACTTACCTACCGCTCAACACC 2305
2041 TTCTCTCTGCGGCGACCTGACACCGGCTCATGCTACACAGAGCTCTGAGCGGCTGTGC 2100
2306 TTCTCTCTGCGGCGACCTGACACCGGCTCATGCTACACAGAGGCTCTGAGCGGCTGTGC 2365
2101 AAACACACCCCGCGAGCCAGCCGAGCTTACAGGACCTGCGGAGCCGCTTTGGCAGAGATC 2160
2366 AAACACACCCCGCGAGCCAGCCGAGCTTACAGGAGCTGCGGAGCCGCTTTGGCAGAGATC 2425
2161 ACAGAGATGTTGACAGCTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATG 2220
2426 ACAGAGATGTTGACAGCTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATG 2485
2221 CACGAACTCAAGAAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2280
2486 CACGAACTCAAGAAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2545
2281 ATCCGCTGCGGCGACCTGACAGAGCTCTGCGGAGAGGAGCTCCAGCAGCGCATGTTCTTC 2340
2546 ATCCGCTGCGGCGACCTGACAGAGCTCTGCGGAGAGGAGCTCCAGCAGCGCATGTTCTTC 2605
2341 CTGTTCAACGACGCTCTGCTATACAGAGCGGAGGAGCTGACGAGCTCTCAATCACTTTAAA 2400
2606 CTGTTCAACGACGCTCTGCTATACAGAGCGGAGGAGCTGACGAGCTCTCAATCACTTTAAA 2665
2401 GTCCAGCGGAGCTCTCCGCTCTATGAGCATGACGATTTGAGAGAGAGAGAGAGAGAGAGAG 2460
2666 GTCCAGCGGAGCTCTCCGCTCTATGAGCATGACGATTTGAGAGAGAGAGAGAGAGAGAGAG 2725
2461 GTGCCCCACCTGCTGACACCTCTCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
2726 GTGCCCCACCTGCTGACACCTCTCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2785
2521 CGGTCCGAGATGAG 2580
2786 CGGTCCGAGATGAG 2845
2581 AGCAGAGAGCCCGCCCTGAGATTCTGAGCAGAGAGCCCGCCCTGACACAGAGTCCCTGAT 2640
2846 AGCAGAGAGCCCGCCCTGAGATTCTGAGCAGAGAGCCCGCCCTGACACAGAGTCCCTGAT 2905
2641 GAAGCCACCGGCTGACAG 2700
2906 GAAGCCACCGGCTGACAG 2965
2701 GAGCGCAGGCGCCGCGACCGGCGACACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2760
2966 GAGCGCAGGCGCCGCGACCGGCGACACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3025
2761 AGCGTCTCCATGTTGAGATTCAAGCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2820
3026 AGCGTCTCCATGTTGAGATTCAAGCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3085
2821 AGGAAATTTAAAAACAGCAACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2880
3086 AGGAAATTTAAAAACAGCAACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3145
2881 CTGTTCTTCTACAAATACACAGAGAGATTTCCCTTTCAGAGCTGCTGCTGCTGCTGCTGCTGCT 2940
3146 CTGTTCTTCTACAAATACACAGAGAGATTTCCCTTTCAGAGCTGCTGCTGCTGCTGCTGCTGCT 3205
2941 TACTGCTCAACATCCCTCTGAGTGTGAGAGAGATTTCCCTTTCAGAGCTGCTGCTGCTGCTGCT 3000
3206 TACTGCTCAACATCCCTCTGAGTGTGAGAGAGATTTCCCTTTCAGAGCTGCTGCTGCTGCTGCT 3265
3001 CACTTCAAGTCCACAGTCTACTTCAAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
3266 CACTTCAAGTCCACAGTCTACTTCAAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3325
3061 ATGGAAGTATCCGAGTGTGACAGAGTGTGCTGCTGCGAGCCCAAGTGTGAGAGAGAGAGAGAG 3120

Db 3326 ATGGAAGTATCCGAGTGTGACAGAGTGTGCTGCTGCGAGCCCAAGTGTGAGAGAGAGAGAGAG 3385
Qy 3121 GAGTCTTGTGTAT 3135
Db 3386 GAGTCTTGTGTAT 3400
RESULT 4
CS130723 5028 bp DNA linear PAT 02-AUG-2005
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 Ornntofc,T.P.
AUTHORS
TITLE
JOURNAL
Aros Applied Biotechnology APS (DX)
FEATURES
Source
1.5028
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="FERM, RhogEF (ARHGEF) and pleckstrin domain pro"
ORIGIN
Query Match 99.9%; Score 3133.4; DB 2; Length 5028;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGAAGTATCCGAGTGTGACAGAGTGTGCTGCTGCGAGCCCAAGTGTGAGAGAGAGAGAGAG 60
Db 337 ATGGAAGTATCCGAGTGTGACAGAGTGTGCTGCTGCGAGCCCAAGTGTGAGAGAGAGAGAGAG 396
Qy 61 GGGATCAGTACCTTGGAGAGTGTGACAG 120
Db 397 GGGATCAGTACCTTGGAGAGTGTGACAG 456
Qy 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 457 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 516
Qy 181 CCGGGAAGTGTGCTGATGAG 240
Db 517 CCGGGAAGTGTGCTGATGAG 576
Qy 241 TTGAGCTCGAGTGTCTGATCAAAAG 300
Db 577 TTGAGCTCGAGTGTCTGATCAAAAG 636
Qy 301 ATTGTAAACAGATTGAAGGCCAAG 360
Db 637 ATTGTAAACAGATTGAAGGCCAAG 696
Qy 361 CCGGCTGACCAACAACTCCAG 420
Db 697 CCGGCTGACCAACAACTCCAG 756
Qy 421 AAGCAGAGCTTGGCTCAAG 480
Db 757 AAGCAGAGCTTGGCTCAAG 816
Qy 481 TCACATTTGTGCAATCTGAGATTTGGGAGTTTGTATGAAGCTTTGAGAGAGAGAGAGAGAGAGAGAG 540
Db 817 TCACATTTGTGCAATCTGAGATTTGGGAGTTTGTATGAAGCTTTGAGAGAGAGAGAGAGAGAGAGAG 876

QY 541 GCAAAAATAATATACATACCTGACGACGACCTAGAGACAAAATCGTGAATTTTCAC 600
| | | | |
Db 877 GCAAAAAATAATATACCTGACGACGACGACCTAGAGACAAAATCGTGAATTTTCAC 936
| | | | |
QY 601 CATTAACCAATTGGACAAACACCAAGCAATCAATTTTCAAGTCTTAGAGATTTGCCGT 660
| | | | |
Db 937 CATTAACCAATTGGACAAACACCAAGCAATCAATTTTCAAGTCTTAGAGATTTGCCGT 996
| | | | |
QY 661 CGGCTAGAGATGATGAAATCGGTTGCAACCGGCGCAAGGACAGGGAAGGACGAAATC 720
| | | | |
Db 997 CGGCTAGAGATGATGAAATCGGTTGCAACCGGCGCAAGGACAGGGAAGGACGAAATC 1056
| | | | |
QY 721 AATCTGCGCTTGGCAACGAGGAATTCATGTTTCAAGGTTTCAATGAATCAATGCC 780
| | | | |
Db 1057 AATCTGCGCTTGGCAACGAGGAATTCATGTTTCAAGGTTTCAATGAATCAATGCC 1116
| | | | |
QY 781 TTCAACTGGGCGCAAGGTGGGAACTGAGCTTCAAGAGAGAGAGGCTTTTCAATCAAGTCTC 840
| | | | |
Db 1117 TTCAACTGGGCGCAAGGTGGGAACTGAGCTTCAAGAGAGAGAGGCTTTTCAATCAAGTCTC 1176
| | | | |
QY 841 CGGCGAGATGCCAATAGTGGTACGAGATACCTTGAAATTCCTGATGGCCAGTCCGAGAT 900
| | | | |
Db 1177 CGGCGAGATGCCAATAGTGGTACGAGATACCTTGAAATTCCTGATGGCCAGTCCGAGAT 1236
| | | | |
QY 901 TTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATCATGCTTCTTTAGACTTTTGA 960
| | | | |
Db 1237 TTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATCATGCTTCTTTAGACTTTTGA 1296
| | | | |
QY 961 GAGCCCAACCAAGCCCAAGCCCGTCTTTAGCCGGGGGTCAATTTTGGTTCAAT 1020
| | | | |
Db 1297 GAGCCCAACCAAGCCCAAGCCCGTCTTTAGCCGGGGGTCAATTTTGGTTCAAT 1356
| | | | |
QY 1021 GGTGGACTCAGAAAGGTTCTGCATATGTTAAAGAGAGACATAAGAAAGGTGACG 1080
| | | | |
Db 1357 GGTGGACTCAGAAAGGTTCTGCATATGTTAAAGAGAGACATAAGAAAGGTGACG 1416
| | | | |
QY 1081 TTTGAAAGAGACAGCAAGATTCATTTATCCGAGGCTTGTTCACAGCTTACAGAA 1140
| | | | |
Db 1417 TTTGAAAGAGACAGCAAGATTCATTTATCCGAGGCTTGTTCACAGCTTACAGAA 1476
| | | | |
QY 1141 CTGAATTCGAAATGCTTGGAGCAGTCTCAGACAGACACAGGCTTACATTTTGAAGAGT 1200
| | | | |
Db 1477 CTGAATTCGAAATGCTTGGAGCAGTCTCAGACAGACACAGGCTTACATTTTGAAGAGT 1536
| | | | |
QY 1201 GCGGAATCTCAGAGGGGGCAGAGCTGCGGCGAGAAAGAAACGAAAGGTTTCCGCGG 1260
| | | | |
Db 1537 GCGGAATCTCAGAGGGGGCAGAGCTGCGGCGAGAAAGAAACGAAAGGTTTCCGCGG 1596
| | | | |
QY 1261 GAGCCGGGGTCCGACCCGAGGCTTGGCCGAGAGAAAGCCCGCGGTTAAACAAGCAGGCG 1320
| | | | |
Db 1597 GAGCCGGGGTCCGACCCGAGGCTTGGCCGAGAGAAAGCCCGCGGTTAAACAAGCAGGCG 1656
| | | | |
QY 1321 GACGAGAGCCGCTCCGGCGCCACGAGAGAAAGAGAGAGAGTCTTAAAGATAGAACCCG 1380
| | | | |
Db 1657 GACGAGAGCCGCTCCGGCGCCACGAGAGAAAGAGAGAGAGTCTTAAAGATAGAACCCG 1716
| | | | |
QY 1381 CAGAATTAACCTCAGGCCCCGAGCCAGACACAGGCTCCCTGATCTGGCAGTCTCACTT 1440
| | | | |
Db 1717 CAGAATTAACCTCAGGCCCCGAGCCAGACACAGGCTCCCTGATCTGGCAGTCTCACTT 1776
| | | | |
QY 1441 TCCGAGCTGTCTGTGAATCTGCAAGGGGAGAGTGGCCCTTGCAACGTTGTCTCC 1500
| | | | |
Db 1777 TCCGAGCTGTCTGTGAATCTGCAAGGGGAGAGTGGCCCTTGCAACGTTGTCTCC 1836
| | | | |
QY 1501 AACCTGAAGCCCCGAGACCAAGAGGCTTCTCCCTTGAATCAGCCCGTGTGTGAATGACG 1560
| | | | |
Db 1837 AACCTGAAGCCCCGAGACCAAGAGGCTTCTCCCTTGAATCAGCCCGTGTGTGAATGACG 1896
| | | | |
QY 1561 GCTTGCCCCCGAGACGAGATGAGATGAGGGCCGAGAGAGAGATTTCCCACTGATAA 1620
| | | | |
Db 1897 GCTTGCCCCCGAGACGAGATGAGATGAGGGCCGAGAGAGAGATTTCCCACTGATAA 1956
| | | | |
QY 1621 GGGTACTTATAGCTAAGGAAGTGTCTACCAACGAGCAACATATCTGAAGATCTCGAA 1680
| | | | |

Db 1957 GGGTACTTATAGCTAAGGAAGTGTCTACCAACGAGCAACATATCTGAAGATCTCGAA 2016
| | | | |
QY 1681 GTTATCACTTGTGGTTTCAAGACCAATGAGCAAAAGAGAGCCGATCCGGAACAACG 1740
| | | | |
Db 2017 GTTATCACTTGTGGTTTCAAGACCAATGAGCAAAAGAGAGCCGATCCGGAACAACG 2076
| | | | |
QY 1741 AAAAGTCAATATTCGCAATTTTGAACCTTTGCAAAATTTTCAATATAATTTTCTCAAG 1800
| | | | |
Db 2077 AAAAGTCAATATTCGCAATTTTGAACCTTTGCAAAATTTTCAATATAATTTTCTCAAG 2136
| | | | |
QY 1801 GAAATTTGACCAACGACTTGCCTGTGGAGAGGCGCTCAATGCCCAATCAGAAATTC 1860
| | | | |
Db 2137 GAAATTTGACCAACGACTTGCCTGTGGAGAGGCGCTCAATGCCCAATCAGAAATTC 2196
| | | | |
QY 1861 CAAGAATCGGAGATGATGATGAGAAACATTCAGGGCATGAAACACTGGCGGCTCAC 1920
| | | | |
Db 2197 CAAGAATCGGAGATGATGATGAGAAACATTCAGGGCATGAAACACTGGCGGCTCAC 2256
| | | | |
QY 1921 CTGTGGAAGCACAGCGAGGCTTGGAGGCGCTGGAATGGAATCAGAGCTCCCGCGG 1980
| | | | |
Db 2257 CTGTGGAAGCACAGCGAGGCTTGGAGGCGCTGGAATGGAATCAGAGCTCCCGCGG 2316
| | | | |
QY 1981 CTGGAATCTTCTGAGAGACTTTGAGCTGCAAGAGGTGTGTACTTACCTGCAACACC 2040
| | | | |
Db 2317 CTGGAATCTTCTGAGAGACTTTGAGCTGCAAGAGGTGTGTACTTACCTGCAACACC 2376
| | | | |
QY 2041 TTCTCTCGCGGCACTGACCGGCTCATGCACTCAAGAGAGTCTGGAAGGCGCTGTGC 2100
| | | | |
Db 2377 TTCTCTCGCGGCACTGACCGGCTCATGCACTCAAGAGAGTCTGGAAGGCGCTGTGC 2436
| | | | |
QY 2101 AAACACACCCGCGCAGGCGACGCGACTTCAGGAGCTGCGAGCGCTTTGGCAGAGATC 2160
| | | | |
Db 2437 AAACACACCCGCGCAGGCGACGCGACTTCAGGAGCTGCGAGCGCTTTGGCAGAGATC 2496
| | | | |
QY 2161 ACGAGATGTGTGACACAGCTCAGGCTCAGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2220
| | | | |
Db 2497 ACGAGATGTGTGACACAGCTCAGGCTCAGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2556
| | | | |
QY 2221 CACGAATCAAGAAATTTTGAATTTGCAATCTTGTGTTTCCGGAAAGGAGTTTC 2280
| | | | |
Db 2557 CACGAATCAAGAAATTTTGAATTTGCAATCTTGTGTTTCCGGAAAGGAGTTTC 2616
| | | | |
QY 2281 ATCCGTCTGGGAGGCTCAGCAAGCTCTCGGGGAAAGGGGCTTCACAGCGCATGTTCTTC 2340
| | | | |
Db 2617 ATCCGTCTGGGAGGCTCAGCAAGCTCTCGGGGAAAGGGGCTTCACAGCGCATGTTCTTC 2676
| | | | |
QY 2341 CTGTTCAACGAGCTCTGCTATACACAGCCGGGGGCTGACGGCTTCCATATAGTTTAA 2400
| | | | |
Db 2677 CTGTTCAACGAGCTCTGCTATACACAGCCGGGGGCTGACGGCTTCCATATAGTTTAA 2736
| | | | |
QY 2401 GTCCAGGGGAGCTCCGCTCTATGGCATGACGATTTGAGAGAGCGAAGACGAGTGGGG 2460
| | | | |
Db 2737 GTCCAGGGGAGCTCCGCTCTATGGCATGACGATTTGAGAGAGCGAAGAGTGGGG 2796
| | | | |
QY 2461 GTGCCCACTGCTGACCCCTCGGGGCGAGCGGCGATCATCTGTGGCGCGCAGTTCT 2520
| | | | |
Db 2797 GTGCCCACTGCTGACCCCTCGGGGCGAGCGGCGATCATCTGTGGCGCGCAGTTCT 2856
| | | | |
QY 2521 CGGTCCGAGATGAGAAATGTGGTTGAGACATCCAGATGGCCATTGACCTGGCGGAAG 2580
| | | | |
Db 2857 CGGTCCGAGATGAGAAATGTGGTTGAGACATCCAGATGGCCATTGACCTGGCGGAAG 2916
| | | | |
QY 2581 AGCAGCAGCCCGGCGCTGAGTCTGTGCGCAGAGCCCGCTGACCAACAGTCCCTGAT 2640
| | | | |
Db 2917 AGCAGCAGCCCGGCGCTGAGTCTGTGCGCAGAGCCCGCTGACCAACAGTCCCTGAT 2976
| | | | |
QY 2641 GAAAGCACCGGGCTGACACAGAGTCAAGAGATGACTGTAGCGCTGCGCATCTGCTG 2700
| | | | |
Db 2977 GAAAGCACCGGGCTGACACAGAGTCAAGAGATGACTGTAGCGCTGCGCATCTGCTG 3036
| | | | |
QY 2701 GAGCCCGAGGCGCCGACCGGGGCAACATTTGTGACGTTGTGTGCAACCGCAACCC 2760
| | | | |

DB 3037 GAGCGCCAGGCCCCCGACCGGCAACACATGTCGCTGTGTCGACCGCAACACC 3096
QY 2761 AGCGTCCCACTGGTGAATTCTCAGCATTCGCGTGGAAATCATGTTCTGTGAAACCTGCTG 2820
DB 3097 AGCGTCCCACTGGTGAATTCTCAGCATTCGCGTGGAAATCATGTTCTGTGAAACCTGCTG 3156
QY 2821 AGGAAATTCAAAACAGCAACGCGGTGGCAGAACTGTGGTGTTCACAAACTTCTGCG 2880
DB 3157 AGGAAATTCAAAACAGCAACGCGGTGGCAGAACTGTGGTGTTCACAAACTTCTGCG 3216
QY 2881 CTGTTCTTCTACAAATACACAGCAACATCATCCCTTCCAGACCTGCTGCTGCGC 2940
DB 3217 CTGTTCTTCTACAAATACACAGCAACATCATCCCTTCCAGACCTGCTGCTGCGC 3276
QY 2941 TACTGCTCAGCAATCCCTCTGTGATCCGAGAAATCCGAAATCATGTTCTGAACTG 3000
DB 3277 TACTGCTCAGCAATCCCTCTGTGATCCGAGAAATCATGTTCTGAACTG 3336
QY 3001 CACTTCAAGTCCCACTGCTCTACTACTTCAAGGCGGAAACGAGTACGTTCAAGGCTG 3060
DB 3337 CACTTCAAGTCCCACTGCTCTACTACTTCAAGGCGGAAACGAGTACGTTCAAGGCTG 3396
QY 3061 ATGGAAGTATCGGACGATGCGACAGCTGTGCTCGGACCCCACTGTTGAGCCACAA 3120
DB 3397 ATGGAAGTATCGGACGATGCGACAGCTGTGCTCGGACCCCACTGTTGAGCCACAA 3456
QY 3121 GAGTCTCTTGTGTAT 3135
DB 3457 GAGTCTCTTGTGTAT 3471

RESULT 5
BD231200 4687 bp DNA linear PAT 17-JUL-2003
LOCUS BD231200 Human cytoskeleton associated proteins.
DEFINITION BD231200
ACCESSION BD231200.1 GI:33040970
VERSION JP 2002526076-A/2.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 4687)
Lal,P., Tang,T.Y., Yue,H., Hillman,J.L., Bandman,O., Corley,N.C.,
Guegler,K.J., Patterson,C., Azimzai,Y. and Baughn,M.R.
Human cytoskeleton associated proteins
Patent: JP 2002526076-A 2 20-AUG-2002;
INCYTE PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002526076-A/2
PD 20-AUG-2002
PR 17-SEP-1999 JP 2000574254
PR 18-SEP-1998 US 60/172226,27-APR-1999 US 60/131321 PI
PREETI LAL, TOM Y TANG, HENRY YUE, JENNIFER L HILLMAN, OLGA PI
BANDMAN
PI NEIL C CORLEY, KARL J GUEGLER, CHANDRA PATTERSON, YALDA AZIMZAI,
PI MARIAH R BAUGHN
PC C12N15/09, A61K38/00, A61K45/00, A61P1/16, A61P3/00, A61P3/06 PC
A61P3/10, A61P5/00,
PC A61P5/38, A61P7/00, A61P7/06, A61P9/00, A61P9/10, A61P11/06 PC
A61P13/00, A61P13/12,
PC A61P17/06, A61P19/02, A61P19/10, A61P21/04, A61P25/00, A61P25/00,
PC A61P25/04,
PC A61P25/14, A61P25/16, A61P25/28, A61P29/00, A61P31/04, A61P31/12,
PC A61P31/18,
PC A61P33/00, A61P35/00, A61P35/02, A61P37/00, A61P43/00, C07K16/47,
PC C07K16/18,
PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12P22/08 PC
C12P1/68, G01N33/15,
PC G01N33/50, G01N33/53, G01N33/53, G01N33/56, C12N15/00, C12N5/00,
PC A61K37/02
CC Incyte ID No: 1320252

PH Key Location/Qualifiers
FT source 1..4687
/organism='Homo sapiens (human)'.
FEATURES
source Location/Qualifiers
1..4687
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 99.9%; Score 3131.8; DB 2; Length 4687;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3133; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGGAAATATAGAGAGAGGCGGACCCAGATATCACTGAGGAGGCCCCGGAAATTTGG 60
DB 90 ATGGGAAATATAGAGAGAGGCGGACCCAGATATCACTGAGGAGGCCCCGGAAATTTGG 149
QY 61 GGGATCAGTACCTTGGAAAGTGGACAGAAAGCCGCCCAACACCTTCAGGAAATCTGCTG 120
DB 150 GGGATCAGTACCTTGGAAAGTGGACAGAAAGCCGCCCAACACCTTCAGGAAATCTGCTG 209
QY 121 TCCATCAAAATCCAGATGCTGATGACACCGAGAGGCAATTTGAAGTTCCACAAAGAGCT 180
DB 210 TCCATCAAAATCCAGATGCTGATGACACCGAGAGGCAATTTGAAGTTCCACAAAGAGCT 269
QY 181 CCTGGGAGAGTGTCTGTGATGCAATTTGCAACCACTCAACTCGTGGAGGTACTAT 240
DB 270 CCTGGGAGAGTGTCTGTGATGCAATTTGCAACCACTCAACTCGTGGAGGTACTAT 329
QY 241 TTTGGCTGAGTTCTCTGATCAAAAAGATCAGGCTGAGTCTCTCTAAACC 300
DB 330 TTTGGCTGAGTTCTCTGATCAAAAAGATCAGGCTGAGTCTCTCTAAACC 389
QY 301 ATTTGAAAACAGATTAGAAAGCCAAAGCACTGTTGTTAGTTGTGTAATTTCTTT 360
DB 390 ATTTGAAAACAGATTAGAAAGCCAAAGCACTGTTGTTAGTTGTGTAATTTCTTT 449
QY 361 CCGCTGACACACACACTCTCCAAAGAACTCACAAGTACTGTTCGCGCTGCAAGCTG 420
DB 450 CCGCTGACACACACACTCTCCAAAGAACTCACAAGTACTGTTCGCGCTGCAAGCTG 509
QY 421 AAGCAGACTTGGCTCAAGGAGGTTGACGTGTAATGACACAGGCGACCTCTTGATT 480
DB 510 AAGCAGACTTGGCTCAAGGAGGTTGACGTGTAATGACACAGGCGACCTCTTGATT 569
QY 481 TCACACTTGTCAATCTGAGATTGGGATTTTGTATGAAGCTTGGACAGAGCACTTA 540
DB 570 TCACACTTGTCAATCTGAGATTGGGATTTTGTATGAAGCTTGGACAGAGCACTTA 629
QY 541 GCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
DB 630 GCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 689
QY 601 CATTAACCAATTGGAACAAACACAGAGATGAGATTTCAGCTCCTAGATTGCGCGT 660
DB 690 CATTAACCAATTGGAACAAACACAGAGATGAGATTTCAGCTCCTAGATTGCGCGT 749
QY 661 CCGCTAGAGATGATGAAATCCGTTTGCACCCGCGCAAGAGAGGAGGAGCAAGATG 720
DB 750 CCGCTAGAGATGATGAAATCCGTTTGCACCCGCGCAAGAGAGGAGGAGGAGGAGATG 809
QY 721 AATTCGCGCTTGGCCCAACCGGAAATTCAGTGTTCAGGGTTTCACTAAGATGATGCG 780
DB 810 AATTCGCGCTTGGCCCAACCGGAAATTCAGTGTTCAGGGTTTCACTAAGATGATGCG 869
QY 781 TTCACCTGGGCGCAAGGTGGGAGGCTGAGCTTCAAGAGAAAGCGCTTCTCAAGCTC 840
DB 870 TTCACCTGGGCGCAAGGTGGGAGGCTGAGCTTCAAGAGAAAGCGCTTCTCAAGCTC 929
QY 841 CCGCAGATGCCAATAGTGCATACAGAGTACCTTGAATTCCTGATGCGCAGTGGGAT 900
DB 930 CCGCAGATGCCAATAGTGCATACAGAGTACCTTGAATTCCTGATGCGCAGTGGGAT 989

QY 901 TTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATCATGCTCTTTAGACTTTTGA 960
| | | | |
Db 990 TTCTGCAATCTCTTCTGAAAAATCTGTGTGAACATCATGCTCTTTAGACTTTTGA 1049
QY 961 GAGCCCAACCAAAAGCCCAAGCCCGTCTCTTTAGCCGGGGGTCAATCATTTGGTTCACT 1020
| | | | |
Db 1050 GAGCCCAACCAAAAGCCCAAGCCCGTCTCTTTAGCCGGGGGTCAATCATTTGGTTCACT 1109
QY 1021 GGTCCGACTCAGAAAGAGTTTCTCGACTATGTATAAGAGAGACATAGAAGGTGACG 1080
| | | | |
Db 1110 GGTCCGACTCAGAAAGAGTTTCTCGACTATGTATAAGAGAGACATAGAAGGTGACG 1169
QY 1081 TTGGAAGAGACACAGCAAGATTCAATTATCCGAGCCTTGCTTCAAGCCTACAGA 1140
| | | | |
Db 1170 TTGGAAGAGACACAGCAAGATTCAATTATCCGAGCCTTGCTTCAAGCCTACAGA 1229
QY 1141 CTGAATTCGAAATGTCTGAGACAGTCTCAGCAGACCAAGCCTTACATTTGAGAAAGT 1200
| | | | |
Db 1230 CTGAATTCGAAATGTCTGAGACAGTCTCAGCAGACCAAGCCTTACATTTGAGAAAGT 1289
QY 1201 GCCGAATCTTCAGAGGGGCCAGAGCTGCGGCGAGAAAGAAACCGAAGTTTCCGCGGG 1260
| | | | |
Db 1290 GCCGAATCTTCAGAGGGGCCAGAGCTGCGGCGAGAAAGAAACCGAAGTTTCCGCGGG 1349
QY 1261 GAGCCGGGGTCCACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTAAACAAGCAGCG 1320
| | | | |
Db 1350 GAGCCGGGGTCCACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTAAACAAGCAGCG 1409
QY 1321 GACGAGCCGCTCTCGGCGCCACGAGAGAAAGAGAGAGTCTTTAAGATAGACCCAG 1380
| | | | |
Db 1410 GACGAGCCGCTCTCGGCGCCACGAGAGAAAGAGAGAGTCTTTAAGATAGACCCAG 1469
QY 1381 CAGAGTAACTCAGGCCCCCGAGCCCAAGACAGAGTCCCTGACTGGAGTCTCAACT 1440
| | | | |
Db 1470 CAGAGTAACTCAGGCCCCCGAGCCCAAGACAGAGTCCCTGACTGGAGTCTCAACT 1529
QY 1441 TCCGAGCTGCTGTGAACTCGAGAGGGGGAGTGGCCCTGCGCAACGTGACTTTGTCTCC 1500
| | | | |
Db 1530 TCCGAGCTGCTGTGAACTCGAGAGGGGGAGTGGCCCTGCGCAACGTGACTTTGTCTCC 1589
QY 1501 AACCTGAGCCCGGACACAGACAGGCTCTCCCTTGATCAGCCCGTCTGATAGACAG 1560
| | | | |
Db 1590 AACCTGAGCCCGGACACAGACAGGCTCTCCCTTGATCAGCCCGTCTGATAGACAG 1649
QY 1561 GCTGTGCCCCGAGCGAGATGAGATGAGGGCCGAGAGAGATTTCCCACTGTATA 1620
| | | | |
Db 1650 GCTGTGCCCCGAGCGAGATGAGATGAGGGCCGAGAGAGATTTCCCACTGTATA 1709
QY 1621 GCGTACTTATAGCTAAGGAAGTGTCTACACCGAGCGAACATATCTGAAGATCTCGAA 1680
| | | | |
Db 1710 GCGTACTTATAGCTAAGGAAGTGTCTACACCGAGCGAACATATCTGAAGATCTCGAA 1769
QY 1681 GTTATATCACTTGGTGTTCAGAGACAGTGAAGCAAGAGAGCCATGCGGAGACATG 1740
| | | | |
Db 1770 GTTATATCACTTGGTGTTCAGAGACAGTGAAGCAAGAGAGCCATGCGGAGACATG 1829
QY 1741 AAAAGTCTATATTCGGAATTTTGAACCTTTGACAAATTTCTAATTTTCTCAAG 1800
| | | | |
Db 1830 AAAAGTCTATATTCGGAATTTTGAACCTTTGACAAATTTCTAATTTTCTCAAG 1889
QY 1801 GAAATTTGACAAAGACTTGCCTGTGAGAGAGCCGCTCAATATGCCAAATCAGAGATTAC 1860
| | | | |
Db 1890 GAAATTTGACAAAGACTTGCCTGTGAGAGAGCCGCTCAATATGCCAAATCAGAGATTAC 1949
QY 1861 CAAAGATAGGAGATGTCTATGCTGAAGAACTTCAAGGACATGAGACCTGGCGGCTCAC 1920
| | | | |
Db 1950 CAAAGATAGGAGATGTCTATGCTGAAGAACTTCAAGGACATGAGACCTGGCGGCTCAC 2009
QY 1921 CTGTGGAAGCACAAGAGGCTTTGAGAGCCTGAGGAATGGAATCAAGAGCTCCGCGGCG 1980
| | | | |
Db 2010 CTGTGGAAGCACAAGAGGCTTTGAGAGCCTGAGGAATGGAATCAAGAGCTCCGCGGCG 2069

QY 1981 CTGGAAGACTTCTGAGAGACTTTGAGCTGAGAGAGTGTGTACTACCGCTCAACACC 2040
| | | | |
Db 2070 CTGGAAGACTTCTGAGAGACTTTGAGCTGAGAGAGTGTGTACTACCGCTCAACACC 2129
QY 2041 TTCTCTCTGCGGCCACTGACACCGGCTCATGTCACTACAGAGGTCTTGAAGGGCTGTGC 2100
| | | | |
Db 2130 TTCTCTCTGCGGCCACTGACACCGGCTCATGTCACTACAGAGGTCTTGAAGGGCTGTGC 2189
QY 2101 AAAACACACCCGCGAGGACGCGGACTTCAAGGAGCTGCCGAGCGCTTTGGCAGAGATC 2160
| | | | |
Db 2190 AAAACACACCCGCGAGGACGCGGACTTCAAGGAGCTGCCGAGCGCTTTGGCAGAGATC 2249
QY 2161 ACGAGATGTGTGCAACAGCTCCACGGTACATGATCAAGATGAGAAATTTCCAGAGCTG 2220
| | | | |
Db 2250 ACGAGATGTGTGCAACAGCTCCACGGTACATGATCAAGATGAGAAATTTCCAGAGCTG 2309
QY 2221 CACGAATCTCAAGAAAGATTTGATTTGGCATTTGATCTTTGTGTTCGGGAAAGGAGTTT 2280
| | | | |
Db 2310 CACGAATCTCAAGAAAGATTTGATTTGGCATTTGATCTTTGTGTTCGGGAAAGGAGTTT 2369
QY 2281 ATCGCTGTGGGAGCTCTGAGCAAGCTCTGGGGAAAGGGGCTTCAGACGCAATGTTCTTC 2340
| | | | |
Db 2370 ATCGCTGTGGGAGCTCTGAGCAAGCTCTGGGGAAAGGGGCTTCAGACGCAATGTTCTTC 2429
QY 2341 CTGTTCAACAGAGTCTGCTATACAGAGCCGGGGGCTGACGGCTTCAATCACTTTTAA 2400
| | | | |
Db 2430 CTGTTCAACAGAGTCTGCTATACAGAGCCGGGGGCTGACGGCTTCAATCACTTTTAA 2489
QY 2401 GTTCCAGGGCAGCTCCCGCTCTATGTGCAATGACATTTGAGAGAGGAAAGCAAGTGGGG 2460
| | | | |
Db 2490 GTTCCAGGGCAGCTCCCGCTCTATGTGCAATGACATTTGAGAGAGGAAAGCAAGTGGGG 2549
QY 2461 GTGCCCACTGCTGACCTCCGAGGGCCAGCGGAGTCCATCATCTGTGGCCGAGTTCT 2520
| | | | |
Db 2550 GTGCCCACTGCTGACCTCCGAGGGCCAGCGGAGTCCATCATCTGTGGCCGAGTTCT 2609
QY 2521 CGGTCCAGATGAGAGATGTGGTTGAGAACATCCAGATGGCCATTGACCTGGCGAGAG 2580
| | | | |
Db 2610 CGGTCCAGATGAGAGATGTGGTTGAGAACATCCAGATGGCCATTGACCTGGCGAGAG 2669
QY 2581 AGCAGACACCCCGCCCTGAGTCTCGGGCCAGAGCCCCCTGACAAAGTCCCTGAT 2640
| | | | |
Db 2670 AGCAGACACCCCGCCCTGAGTCTCGGGCCAGAGCCCCCTGACAAAGTCCCTGAT 2729
QY 2641 GAAAGCACCGCGGCTGACACAGAGATCAGAGATGACTGAGCGCCTCGGCACATCGCTG 2700
| | | | |
Db 2730 GAAAGCACCGCGGCTGACACAGAGATCAGAGATGACTGAGCGCCTCGGCACATCGCTG 2789
QY 2701 GAGCGCCAGGCCCGCACCGCGGCAACAATGTGTGACAGTGTGTGCAACCGCAACC 2760
| | | | |
Db 2790 GAGCGCCAGGCCCGCACCGCGGCAACAATGTGTGACAGTGTGTGCAACCGCAACC 2849
QY 2761 AGCGTCTCATATGTGGAATTTGACATGCAAGTGAAGATTCAGTTGTCTGAAACCTGCTG 2820
| | | | |
Db 2850 AGCGTCTCATATGTGGAATTTGACATGCAAGTGAAGATTCAGTTGTCTGAAACCTGCTG 2909
QY 2821 AGGAATTTCAAAAACAGCAAGGTTGAGAGAGCTGTGGTGTGTTCACAACTTTCTGC 2880
| | | | |
Db 2910 AGGAATTTCAAAAACAGCAAGGTTGAGAGAGCTGTGGTGTGTTCACAACTTTCTGC 2969
QY 2881 CTGTTCTTCTACAAATCAACACAGACATATCCCTTGCCAGCTGTGCTCGGC 2940
| | | | |
Db 2970 CTGTTCTTCTACAAATCAACACAGACATATCCCTTGCCAGCTGTGCTCGGC 3029
QY 2941 TACTGCTCAACATCCCTCTGAGTCCGAGAACATCCAGAAAGATAGTGTTCAGAGCTG 3000
| | | | |
Db 3030 TACTGCTCAACATCCCTCTGAGTCCGAGAACATCCAGAAAGATAGTGTTCAGAGCTG 3089
QY 3001 CACTTCAAGTCCACAGTCTACTTCAAGGCGGAAACGAGTACAGTTTCGAAAGGTGG 3060
| | | | |
Db 3090 CACTTCAAGTCCACAGTCTACTTCAAGGCGGAAACGAGTACAGTTTCGAAAGGTGG 3149
QY 3061 ATGGAAGTATCCGAGTGGCCACAGCTTGTGCGGACCCCAAGTGTGAGCCACAA 3120

Db 3150 ATGGAAGTATCCGACAGTCCGCGGACCCAGCCAGTGTGAGTCACAAA 3209
Qy 3121 GAGTCTCTTGTGTAT 3135
Db 3210 GAGTCTCTTGTGTAT 3224

RESULT 6
BC071592 4932 bp mRNA linear PRI 28-JUL-2005
LOCUS Homo sapiens FERM, RhogEF (ARHGEF) and pleckstrin domain protein 1
DEFINITION (chondrocyte-derived), mRNA (cDNA clone MGC:87400 IMAGE:30342239),
complete cds.
BC071592
BC071592.1 GI:47940451
VERSION MGC.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 4932)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haide, F.,
Diatchenko, U., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schecterson, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carneiro, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramsen, R.D., Mulvihy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4932)
NIH MGC Project
Direct Submission
Submitted (01-JUN-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-ehgc.stanford.edu>
Contact: (Dickson, Mark) mc@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
Series: IRAP Plate: 168 Row: 1 Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 48928036.
Location/Qualifiers
1. .4932

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:87400 IMAGE:30342239"
/issue_type="Placenta, normal"
/clone_lib="NIH MGC_147"
/lab_host="DH10B"
/note="Vector: pBluescriptCR"
1. .4932
/gene="FARP1"
/note="synonyms: CDEP, MGC87400, PLEKHG2"
/db_xref="GeneID:10160"
/db_xref="MIM:602654"
240_ .3470
/gene="FARP1"
/codon_start=1
/product="FARP1 protein"
/protein_id="AAH71592.1"
/db_xref="GI:47940452"
/db_xref="GeneID:10160"
/db_xref="MIM:602654"
/translation="MGIEQRPPPSRLAPENSGISTLERGGKPPPTSGKVLTKI
OMLDDYQAEVFORAPGVLDVAVCNHLNVEGDFGLFPPDHKKITWLDLKLIV
KOIRPKHVVVVFVVFPPPDHQLQOEHLTYLFAIQVQDLQGRITCNDSALLI
SHVQSEIDGPDALDRHLAKKVIPOODALDKTIVEFHNNHIGOTPAESDQLLEI
ARLEWYGRILHAKDRGCKINLAVANTGIIIVFOGRTKINFMWAKVKRLSPKRRP
LIKRPDANSADVDPTLEFLMASRDRPKSWIKCEVHRAFRPLPEEPKPKPVLPBRS
SSFRSGRQKQKQVLDVKEGHNKQVPEKHSKHSIRSLASQTEPLNSEVLESOQS
TSITLFBEGASRPGSCRGKPKVASGPRSPAPRSPAGNTPANVTLSPLNISPDKO
EEVVVDNRQKSKPOPPQSGTSLNPSHLSVVSOGVAPANTVLSPLNISPDKO
ASPLISPLNDQACPTDDEGRKRPPTDRAVFAKESTTERETVLYKDLVITSMF
QSVSKEDNPMALXSLITPNPPLHFKHTNPLKETEORLALMEGSMNOIRPYQIG
DMLKRTQKMKHLAHLIKYSBALXLENGISRSRLNFCRDFELQKCYLPINPL
LPLRHLKQVLEKLCNHPSPADFPDCCALKEITEMAQLHGTWKENFQKL
HEHKDLQIDIMLVVGRGSPSLKMTPPHAGARIPCAPERRPLLVKPEIRLGLSL
KLSGKQSRMFFLFNDVLLYSRGLTASNOFVHQLPLVGTITSESDENGVPDLT
TSGKQOSIIVASRSRMEKWEVDIOMLIDLAEKSSPAPELTASPDNKSDEAT
ADQSESDILASRSLERQAPRGNTWVHCMTHTVSQVLDFTAVNOLSGNLLR
KXNSGMOKLVAVPTNPLCPFKYSQDNHPLASPLVLTYSLTIPSESENIOQDYVFK
LHFKSHVYTFRASEYTFERMEVIRKATSSASRPVLSHKSPLYT"
ORIGIN
Query Match 96.6%; Score 3028.8; DB 5; Length 4932;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 3133; Conservative 0; Mismatches 2; Indels 93; Gaps 1;
Qy 1 ATGGAGAATATAGACAGAGCCGACCCAGATCAAGCTGGGGGCCCCGGAATTCG 60
Db 240 ATGGAGAATATAGACAGAGCCGACCCAGATCAAGCTGGGGGCCCCGGAATTCG 239
Qy 61 GGGATCAGTACCTTGGAAAGTGGACGAAGCCGCCCAACACCTTCAGGAAAATCTCG 120
Db 300 GGGATCAGTACCTTGGAAAGTGGACGAAGCCGCCCAACACCTTCAGGAAAATCTCG 359
Qy 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGGACATTTGAAGTTCACAAAGAGCT 180
Db 360 TCCATCAAAATCCAGATGCTGATGACACCCAGAGGACATTTGAAGTTCACAAAGAGCT 419
Qy 181 CTTGGAGAGTGTCTGTGATGACAGTTTCAACACCTCAACTCTGTGAAGGTGACTAT 240
Db 420 CTTGGAGAGTGTCTGTGATGACAGTTTCAACACCTCAACTCTGTGAAGGTGACTAT 479
Qy 241 TTTGGCTCTGAGTTTCTGTATCAAAAAGATACGGTGTGGTATCTCTTAAACCC 300
Db 480 TTTGGCTCTGAGTTTCTGTATCAAAAAGATACGGTGTGGTATCTCTTAAACCC 539
Qy 301 ATTGTAAACAGATTGAAGGCCAAAGCAAGTGTGTGAAGTTTGTGTAATTCCTT 360
Db 540 ATTGTAAACAGATTGAAGGCCAAAGCAAGTGTGTGAAGTTTGTGTAATTCCTT 599
Qy 361 CCGGCTGACCAACAGACTCCAAAGAACTCACAAGGTACTGTTCGCGCTGACGGTG 420
Db 600 CCGGCTGACCAACAGACTCCAAAGAACTCACAAGGTACTGTTCGCGCTGACGGTG 659

OY 421 AAGCAGAGCTTGGCTCAAGGCAAGTGAAGTGAATGACACAGCCAGCTCTTGATT 480
DB 660 AAGCAGAGCTTGGCTCAAGGCAAGTGAAGTGAATGACACAGCCAGCTCTTGATT 719
OY 481 TCACACATTTGTGCAATCTGAGATTTGGGAAATTTGATGAAGCTTGGACAGAGACCTTA 540
DB 720 TCACACATTTGTGCAATCTGAGATTTGGGAAATTTGATGAAGCTTGGACAGAGACCTTA 779
OY 541 GGAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
DB 780 GGAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 839
OY 601 CATTAACCATTTGAGCAAAACACAGAGCAATCAGATTTCCAGCTCTTAGAGATTTGCCGT 660
DB 840 CATTAACCATTTGAGCAAAACACAGAGCAATCAGATTTCCAGCTCTTAGAGATTTGCCGT 899
OY 661 CGGCTAGAGATGTATGAAATCCGGTTGCAACCCGGCCAAAGACAGGGAAGGCAAGAAATC 720
DB 900 CGGCTAGAGATGTATGAAATCCGGTTGCAACCCGGCCAAAGACAGGGAAGGCAAGAAATC 959
OY 721 AATTCGGCCGTGGCCCAACAGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGACC 780
DB 960 AATTCGGCCGTGGCCCAACAGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGACC 1019
OY 781 TTCAACTGGGCGCAAGGTGCGGAAAGTGAAGCTTCAAGAGAAAGGCTTTCATCAAGCTC 840
DB 1020 TTCAACTGGGCGCAAGGTGCGGAAAGTGAAGCTTCAAGAGAAAGGCTTTCATCAAGCTC 1079
OY 841 CGGCGAGATGCCAATAGTGGTACAGAGATACCTTGAAATTCCTGATGGCCAGTCCGGAT 900
DB 1080 CGGCGAGATGCCAATAGTGGTACAGAGATACCTTGAAATTCCTGATGGCCAGTCCGGAT 1139
OY 901 TTCTGAAATGCTTTCGAAAAATCTGTGTGAACATCAGCTCTTGAAGCTTTTGA 960
DB 1140 TTCTGAAATGCTTTCGAAAAATCTGTGTGAACATCAGCTCTTGAAGCTTTTGA 1199
OY 961 GAGCCCAACCAAAAGCCCAAGCCCGTCTCTTAAAGCGGGGTCATCATTTGCGTTCAGT 1020
DB 1200 GAGCCCAACCAAAAGCCCAAGCCCGTCTCTTAAAGCGGGGTCATCATTTGCGTTCAGT 1259
OY 1021 GGTGCGACTCAGAGAGAGTTCCTGACTATGTTAAAGAGAGAGACATTAAGAGGTGACG 1080
DB 1260 GGTGCGACTCAGAGAGAGTTCCTGACTATGTTAAAGAGAGAGACATTAAGAGGTGACG 1319
OY 1081 TTGGAAGAGAGAGAGAGATTCATTCATTCGAGGCTTGCCTTACAGGCTTACAGAA 1140
DB 1320 TTGGAAGAGAGAGAGAGATTCATTCATTCGAGGCTTGCCTTACAGGCTTACAGAA 1379
OY 1141 CTGAATTCGGAAGTGTGAGAGAGTCTCAGACAGAGACAGCCTTACATTTGAGAGAGT 1200
DB 1380 CTGAATTCGGAAGTGTGAGAGAGTCTCAGACAGAGACAGCCTTACATTTGAGAGAGT 1439
OY 1201 GCGGAATTCACAGGGGCGCAAGAGCTCCGCGAGAGAAAGAACCGAAGTTTCCGCGG 1260
DB 1440 GCGGAATTCACAGGGGCGCAAGAGCTCCGCGAGAGAAAGAACCGAAGTTTCCGCGG 1499
OY 1261 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAGAGCCCGCGGGTAAACAAGCAGG 1320
DB 1500 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAGAGCCCGCGGGTAAACAAGCAGG 1559
OY 1321 GACGAGCCGCTCGCGCCCAAGAGAGAGAGAGAGAGTCTGTTAAGATAGAACCCAG 1380
DB 1560 GACGAGCCGCTCGCGCCCAAGAGAGAGAGAGAGAGTCTGTTAAGATAGAACCCAG 1619
OY 1381 CAGAGTAACTCTAGCCCCCGAGCCCAAGACAGAGTCCCTGACTGCGAGTCTCACTT 1440
DB 1620 CAGAGTAACTCTAGCCCCCGAGCCCAAGACAGAGTCCCTGACTGCGAGTCTCACTT 1679
OY 1441 TCCGAGCTGTCTGTGAATCTGACAGGGGGAGTGGCCCTGCGCAAGTGAACCTTGTCTCC 1500
DB 1680 TCCGAGCTGTCTGTGAATCTGACAGGGGGAGTGGCCCTGCGCAAGTGAACCTTGTCTCC 1739

OY 1501 AACCTGAGCCCGACACCAAGCAGGCTCTCCCTTGATCAGCCGCTGTGATGACAG 1560
DB 1740 AACCTGAGCCCGACACCAAGCAGGCTCTCCCTTGATCAGCCGCTGTGATGACAG 1799
OY 1561 GCGTCCCGCCCGAGCGAGATGAGAGTGGGCGGAGAGAGATTTCCCACTGATAA 1620
DB 1800 GCGTCCCGCCCGAGCGAGATGAGAGTGGGCGGAGAGAGATTTCCCACTGATAA 1859
OY 1621 GGTACTTCATGACTAAGAGTGTCTACACAGAGCGAATCATCTGAGAGATCTGAA 1680
DB 1860 GGTACTTCATGACTAAGAGTGTCTACACAGAGCGAATCATCTGAGAGATCTGAA 1919
OY 1681 GTTATCACTCTGTGTTTCAAGACAGTGAACAAAGAGAGAGCCATGCGGAGCACTG 1740
DB 1920 GTTATCACTCTGTGTTTCAAGACAGTGAACAAAGAGAGAGCCATGCGGAGCACTG 1979
OY 1741 AAAAGTCTCATATTTCCGAAATTTTGAACCTTTGGCAAAATTTCAATTAATTTTCTCAG 1800
DB 1980 AAAAGTCTCATATTTCCGAAATTTTGAACCTTTGGCAAAATTTCAATTAATTTTCTCAG 2039
OY 1801 GAAATTTGAGCAACCTTGCCCTGTGGAGAGCCGCTCAATGCCCCAAATCAGAGATTAC 1860
DB 2040 GAAATTTGAGCAACCTTGCCCTGTGGAGAGCCGCTCAATGCCCCAAATCAGAGATTAC 2099
OY 1861 CAAGAATTCGGGATGTATGCTGAAGAACATTCAGGGCATGAGCACCTGGCGCTCAC 1920
DB 2100 CAAGAATTCGGGATGTATGCTGAAGAACATTCAGGGCATGAGCACCTGGCGCTCAC 2159
OY 1921 CTGTGGAAGACAGAGAGGCTTTGAGAGCCCTTGGAGAAATGGAATGAAGCTCCGGCG 1980
DB 2160 CTGTGGAAGTAAAGAGAGGCTTTGAGAGCCCTTGGAGAAATGGAATGAAGCTCCGGCG 2219
OY 1981 CTGGAAGACTTCTGACAGACTTTGAGCTGCAAGAGTGTGTTACTACCGCTCAACACC 2040
DB 2220 CTGGAAGACTTCTGACAGACTTTGAGCTGCAAGAGTGTGTTACTACCGCTCAACACC 2279
OY 2041 TTCTCTCTGCGGCACTGACCGGCTCATGCACTAACAGAGTCTTGAGCGGCTGTGC 2100
DB 2280 TTCTCTCTGCGGCACTGACCGGCTCATGCACTAACAGAGTCTTGAGCGGCTGTGC 2339
OY 2101 AAAACACACCGCCGAGCCAGCCGACCTTCAGAGGATCTGCCGAGCCGCTTTGGCAGAGATC 2160
DB 2340 AAAACACACCGCCGAGCCAGCCGACCTTCAGAGGATCTGCCGAGCCGCTTTGGCAGAGATC 2399
OY 2161 ACGAGATGTGTGACAGCTCCACGCTACAGATGATCAAGATGAGAAATTTCCAGAGCTG 2220
DB 2400 ACGAGATGTGTGACAGCTCCACGCTACAGATGATCAAGATGAGAAATTTCCAGAGCTG 2459
OY 2221 CACGAATCTCAAGAAAGATTTGATGGAATGACAAATCTTGTGTTCCGGGAAG----- 2273
DB 2460 CACGAATCTCAAGAAAGATTTGATGGAATGACAAATCTTGTGTTCCGGGAAGCGCTGGC 2519
OY 2274 ----- 2273
DB 2520 TCCTTTTCCCTTATGAGGATTCACACTTAAGACAAAGCCAGGCAATCCATGTGCTCA 2579
OY 2274 -----GGAATTCATCCGCTTGGGAGGCTCAGAAAGCTC 2307
DB 2580 GAGAGAGAGCCATTGTATTAAGAGATTCACCGTCTGTGGAGAGCTCAGAAAGCTC 2639
OY 2308 TCGGGGAAGGGGCTCCAGAGCGCATGTTCTTCTGTTCACAGAGCTCTGCTATACAG 2367
DB 2640 TCGGGGAAGGGGCTCCAGAGCGCATGTTCTTCTGTTCACAGAGCTCTGCTATACAG 2699
OY 2368 AGCCGGGGGCTGACGGCTTCCAAATGATTTAAAGTCCAGGAGCACTCCCGTCTATAGG 2427
DB 2700 AGCCGGGGGCTGACGGCTTCCAAATGATTTAAAGTCCAGGAGCACTCCCGTCTATAGG 2759
OY 2428 ATGAGATTTGAGAGAGAGAGACAGAGTGGGGGAGGCCCACTGCGCTGACCTCCGGGGC 2487
DB 2760 ATGAGATTTGAGAGAGAGAGAGAGAGTGGGGGAGGCCCACTGCGCTGACCTCCGGGGC 2819
OY 2488 CAGCGGAGTCCATCATCTGTGGCGCACAGTTCTCGGTCCGAGATGAGAAAGTGGGTTGAG 2547

Accession	Version	Keywords	Source	Organism	Reference	Authors
Db						
2820						
Qy						
2548						
Db						
2880						
Qy						
2608						
Db						
2940						
Qy						
2668						
Db						
3000						
Qy						
2728						
Db						
3060						
Qy						
2788						
Db						
3120						
Qy						
2848						
Db						
3180						
Qy						
2908						
Db						
3240						
Qy						
2968						
Db						
3300						
Qy						
3028						
Db						
3360						
Qy						
3088						
Db						
3420						
RESULT 7						
BC030329						
LOCUS						
DEFINITION						
ACCESSION						
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						

TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS	JOURNAL	REMARK	COMMENT
Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimmold, J., Schmutz, J., Myers, R. M., Butcherfield, Y. S., Krzywinski, M. I., Skalska, U., Smalins, D. E., Schercher, A., Schein, J. E., Jones, S. J. and Marra, M. A.							
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	Proc. Natl. Acad. Sci. U.S.A.	99 (26), 16959-16903 (2002)					
12477932							
2 (bases 1 to 3902)							
Strauberg, R.							
Direct Submission							
Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA							
NIH-MGC Project URL: http://mgc.nci.nih.gov							
Contact: MGC help desk							
Email: gcgaps-remail.nih.gov							
Tissue Procurement: Gilbert Smith, Ph.D.							
cDNA Library Preparation: Life Technologies, Inc.							
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILML)							
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;							
Web site: http://www.nisc.nih.gov/							
Contact: nisc_mgc@nrgi.nih.gov							
Akter, N., Ayele, K., Beckstrom-Sternberg, S. M., Benjamin, B., Blakesley, R. W., Bouffard, G. G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N. L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maburo, Q. L., Masiello, C., Maiberti, B., Mastrian, S. D., McCloskey, J. C., McDowell, J., Pearson, R., Stanciripob, S., Thomas, P. J., Touchman, J. W., Young, A., Vogt, J. L., Walker, M. A., Wetherby, K. D., Wiggins, L., Young, A., Zhang, L. H. and Green, E. D.							
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILML at: http://image.llnl.gov							
Series: IRAX Plate: 65 Row: m Column: 19							
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.							
Location/Qualifiers							
1. .3902							
/organism="Mus musculus"							
/mol_type="mRNA"							
/strain="FVB/N-3"							
/db_xref="taxon:10090"							
/clone="IMAGE:5376197"							
/tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."							
/clone_1lb="NCI CGAP_Mam2"							
/lab_host="DH10B"							
/note="vector: pCMV-Sport6"							
<1. .3902							
/gene="Farp1"							
/note="synonym: Cdep"							
/db_xref="GeneID:223254"							
/db_xref="MGI:2446173"							
<1. .2510							
/gene="Farp1"							
/codon_start=3							
/product="Farp1 protein"							
/protein_id="AAH30329.1"							
/db_xref="GI:20987936"							
/db_xref="GeneID:223254"							
/db_xref="MGI:2446173"							
/translation="TRPVARRLRLEWYGIIRLHPADREGTKINLAIVANTGIIIVFGGFTKINFMKAVKRLSPKRFILKRPDYNSSVODLFLFMGRDCKSPFKVCVHNHFLPFEERKPKRPVLFSGSGSPFGSGTQKQVLDYVGECHKVQFPRKISKVSHNLFVSPQTPNENVEPKQSPASISLTGEGTESPGQSCCOAKETKACTLELGHOSPALEKSGPAAADGTTVYPPEREEREGGKDIKPSNPORPDSSTSLTSPSHLSISINSKGGAPPAVTLSPVLSKDNKQASFLPILANDACPTDDEEGRRRRFTPDKAYIYIKVSTERTYTKDLDELVAISWQSTVSKEDSPKALSKLIPNFPDLHKFHNPLKEIE							

ORLALMEGRSNNAHVBDYORIGDVMILKNIOGMKHLAHLMKHSEALELETISIKSRR
LEHPCRDPELOKCYCLPLNTPLRLPLRMLHYKHYLERCKKHPRNHADFRDCRALA
ETTEMVAOLHGMIMKIMENFOKLHELKDLIGDNLVTPEBEPTRLGISKISGKGLQO
RMPPLPNDLYLTSRGLTASNOFKVHGOCLPYGMTEBSEEWGPHCLTLRAGOSI
IYAASSRSEMEKEMEDIQNALIDAKSNFTPEBLASSPPDKSPDEATADQSEBD
LSASRSLRQAHRNGNTVHVCWHRSTSVSWDPSIAVENLSGLNLRKFNNSGWO
KLAVVETNCLPFYKSHODSHPLA.PLIGYSLTTPSEENIHKOYVFLHFKSHVYV
FASSEYTERFERMEVIRSATSSASRAHILSHKESHLX"
999. 1556

misc_feature
/gene="Farpi"
/note="RhoGEF; Region: Guanine nucleotide exchange factor
for Rho/Rac/Cdc42-like GTPases"
/db_xref="CD:smart00325"
1695. 1946
/gene="Farpi"
/note="PH; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntenophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinaemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CD:smart00233"
2175. 2465

misc_feature

/gene="Farpi"
/note="PH; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntenophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinaemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CD:smart00233"

ORIGIN

Query Match 61.1%; Score 1916.8; DB 6; Length 3902;
Best Local Similarity 86.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 342; Indels 9; Gaps 2;

QY 642 GCTCCTAGAGATTGCCGCTCGCTAGAGATGTATGAAATCCGGTTGCAACCGGCAAGGA 701
DB 5 GCGTCCGAGAGTGCGCCGTGCTAGAGATGTATGAAATCCGGTTGCAACCGGCAAGGA 64
QY 702 CAGGGAAGCAGAGATCAATCTGCGCTGCGCAACGCGAATCTGTGTTTCAAGG 761
DB 65 CAGGGAAGCAGAGATCAATCTGCGCTGCGCAACGCGAATCTGTGTTTCAAGG 124
QY 762 TTTCAGTAAATGCTTCAACTGCGGCAAGGCGGGAAGCTGAGCTTCAAGAGAA 821
DB 125 CTTCAGAAAGATCATCTTTTAACTGGGCAAGGTCGGAAGCTGAGCTTCAAGGAA 184
QY 822 GCGCTTTCTCATAGCTCCGCGCAAGATCCAAATAGTGCATACAGATACCTTGAAT 881
DB 185 GCGCTTTCTTATCAAGCTCAAGCTGACGTGAATACCTATACAGAGAACCTTGAAGT 244
QY 882 CCTGATGCGCAAGTCTGCGAATCTTCTGGAATACTGTGTGAACATCATG 941
DB 245 CCTGATGCGCGGCGGAGCTTCTGCAAGTCTTCTGGAATAATGTGTGAGAGCACTATC 304
QY 942 CTTCTTAGACTTTTGAAGAGCCAAAGCCCAAGCCGCGCTCTTTAGCCGGG 1001
DB 305 CTTCTTAGACTTTCTGAAGAACTTAAGCCAAAGCCGCTTCTCTTCAAGCCAGG 364

QY 1002 GTCATCATTTCCGTTCACTGATGCTCGACTCAGAAAGCGTTCTCGACTATGTTAAAGAG 1061
DB 365 GTTCGCTCTTCGGTTTAAAGGTCGAGCTCAGAAAGCAAGTTCTGATTAATGTTAAAGAG 424
QY 1062 AGGACATTAAGAAAGTGCAGTTTGAAGAGACAGCAAGATTCATTCTATCCGAGCT 1121
DB 425 CGGCAACAAGAAAGGTCAGTTTGAAGAAAGACAGCAAGATTCCTCCACAGAGAGCT 484
QY 1122 TGCTTCAAGCGCTCAGAACTGAATTCGAAAGTGTGAGAGCAGTCTCAGACAGACAG 1181
DB 485 GGTCTCAGAGCCACAGACAGCAAGATTCAGAAAGTGCAGAAAGTCTCAGACAGAGCGA 544
QY 1182 CCTTACATTTGAGAAAGTGCCGAATCTCCAGGGGCGAAGCTCCGCGAGAGAAAGGA 1241
DB 545 CTTTACGTTTGAAGAAAGTGAACGAGATCTCCGGGTGGCCAGAGCTCCAGCAAGCGAAGGA 604
QY 1242 ACCGAAAGTTTCCGCGGGGAGACCGGGGTCCGACCCGAGCCCTGGCGCGAAGAAAGCC 1301
DB 605 AACAAAGCTGTATCCTTGGAGTTGGACCTCACCAAGAGCCCGGCTCTTAAGAGCTC 664
QY 1302 CGCGGTTAAAGAGCGAGCGAGCGGCGCTCGCGCCACAGAGAAAGAGAGAGT 1361
DB 665 TCCAGAGAGCAAGAGCTGCAGATGGAACCAACCTGTGTCCGCTGAGAGAAAGAGAGGA 724
QY 1362 CGTT-----AAGATTAAGAACCCAGCAGAGTAAACCTGAGCCCGCAGCCAAAGCAG 1415
DB 725 AAGAGGTGCAAGAGACGCGCATCTCGGCAAGTAACTCAGCCCGCAGCCAAAGCAGG 784
QY 1416 CTCCCTGAGTGGACATCTCTCACTTTCCGAGCTGTCTGTGAACCTGCGAGGGGAGTGGC 1475
DB 785 CTCCCTGAGTGGTACCTCTCACTTTCAAGAGTGTCTCACTCACTCAAGGGAGAGCGC 844
QY 1476 CCTGCGCAAGCGTACCTTGTCTCCCACTGAGCCCGCAGCAAGAGAGGCTCTCCCT 1535
DB 845 CCCAGCAAGCGTACCTGTCTCTCCCACTGAGCTCTCACTCACTCAAGAGAGGCTCTCCCT 904
QY 1536 GATCAGCCCGCTGCTGAATGAACAGAGCTGCCCCGAGCGAGCAGATGAAGATGAGGCGC 1595
DB 905 GATCAGCCCGCTGCTGAATGAACAGAGCTGCCCCGAGCGAGCAGATGAAGATGAGGCGC 964
QY 1596 GAGGAAGATTTCCCACTGATTAAGCGTACTTCACTTAAAGAGAGTCTAACACCGA 1655
DB 965 GCGAAAGAGATTTCCCACTGATTAAGCGTACTTCACTTAAAGAGAGTCTAACACCGA 1024
QY 1656 GCGAATATCTGAAGATCTCGAAGTATCACTTGTGTTTCAAGACAGTGAAGCA 1715
DB 1025 GCGCAGTACCTGAAGAGACCTTGAAGTCACTTGTGTTTCAAGACAGTGAAGCA 1084
QY 1716 AGAGGACGCTATGCGGAGACCTGAAGATCTCATATTTCCGAATTTTGAACCTTTGCA 1775
DB 1085 AGAGGACGCTATGCGGAGACCTTGAAGATCTCATATTTCCGAATTTTGAACCTTTGCA 1144
QY 1776 CAAATTTCACTATTTTCTCAAGGAATTAAGAGAGAGCTTGCCTGTGGGAAGGCGC 1835
DB 1145 CAAATTTCACTATTTTCTCAAGGAATTAAGAGAGAGCTTGCCTGTGGGAAGGCGC 1204
QY 1836 CTCAATGCCCCCAATCA---GAGATTACCAAGAAATCGCGATGTCATGCTGAAGAACT 1892
DB 1205 CTCAATGCCCCCAATGCTCAAGAGAGATTAACAAAGATCGGGGACGTCATGCTGAAGAACT 1264
QY 1893 TCAGGGCATGAAGCACTGGGGGCTCACTGTGGAAGCAAGCGAGGCTTGAAGGCTT 1952
DB 1265 TCAGGGCATGAAGCACTGGGGGCTCACTGTGGAAGCAAGCGAGGCTTGAAGGCTT 1324
QY 1953 GGAAGATGAAGATCAAGAGCTCCGCGGCTGGAAGACTTTCGCAAGAGCTTGAAGTGA 2012
DB 1325 GGAAGATGAAGATCAAGAGCTCCGCGGCTGGAAGACTTTCGCAAGAGCTTGAAGTGA 1384
QY 2013 GAAAGTGTGTTTAACTTACAGGCTCAACACTTCTCTGTGGGCGCACTGACCGGCTATGCA 2072
DB 1385 GAAAGTGTGTTTAACTTACAGGCTCAACACTTCTCTGTGGGCGCGCTGACCGGCTATGCA 1444
QY 2073 CTAACAAGAGGTTCTTGAAGCGGCTGTGCAAAACCAACCGCGGAGCCAGCAGCTTAG 2132

Db	1445	CTATTAGAGATTCTCTGGAGAGGCTGTGCAGACCAACCCACCAACACGCCGCACTTACG	1504
QY	2133	GGACTGCCGAGCCGCTTTTGGACAGATCACGGAGATGGTGGCAGACTCCACGGTACGAT	2192
Db	1505	GGACTGCAGACTGCGCTGCGCGAGATCACAGAGATGGTGGCAGCTGCATGGATACAT	1566
QY	2193	GATCAAGATGGAGAAATTTCCAGAGCTGCAGAACTCAAGAAAGATTTGATTGGCATTTGA	2255
Db	1565	GATCAAGATGGAGAACTTCCAGAGCTGCAGAAAGTCAAGAAAGATTCGATTGGCATTTGA	1624
QY	2253	CAATCTTGTGGTTCCGGGAGAGGAGTTATCCGTCGTGGGCAAGCCCTCAGCAAGCTCTCGGG	2312
Db	1625	CAATCTTGTGACCCAGAGAGGAGCTTATCCGCTGGGCAAGCCCTCAGCAAGCTCTCCGG	1688
QY	2313	GAAAGGGGCTCCAGCAGCCGATGTTCTTCTGTTCACACAGCTCTGTCTATACACGAGCG	2372
Db	1685	GAAAGGGGCTCAGCGAGCGCATGTTCTTCTGTTCACAGATGTTCTGTATTACACGAGCG	1744
QY	2373	GGGGCTGACGGCCCTCCATCACTGTTTAAAGTCACGGGCAGCTCCCGCTCTATTGGCATAC	2432
Db	1745	GGGACTGACGGCATCTATCAAGTTTAAAGTCAACGGACAGCTCCCATCTATTGGCATAC	1804
QY	2433	GATTGAGAGAGAGCGMAAGCAGATGGGGGGGTGCCCACTGCGCTGACCCCTCCGGGGCAGCG	2492
Db	1805	GATTGAGAGAGAGTAAAGAGAGTGGGGGTGTGCCCACTGCGCTGACCCCTCCGGGGCAGCG	1864
QY	2493	GCACTGCATCATCGTGGCCGCGCAGTCTCGGTCCGAGATGGAGAGTGGGTTGAGACAT	2552
Db	1865	GCAATCTATCATCGTGGCTGCGCAGTCCCGGTCAAGAGATGGAAAAATGAGATGGAGACAT	1924
QY	2553	CCAGATGGCCATTGACTCGCGCGAGAGAGACAGACGCCGCCCTCTGATTTCTTGGCCAG	2612
Db	1925	CCAGATGGCCATTGACTCGCGAGAAAAAGAAATGAAACCCACCCCGAGCTCTTGGCCAG	1984
QY	2613	CAGCCCCCTTGAACAACAAGTCCCGTGAATGAAGCCACCGGGCTGACAGAGTACAGAGA	2672
Db	1985	CAGCCCCCTTGAACAACAAGTCCCGGATGAAGCCACAGCGGTGACAGAGTCAAGAGA	2044
QY	2673	TGACTTGAACGCGCTTCGCGCAGCATCGCTGGACGCGCAGGCCCGCAGCAGCGGCAACAAT	2732
Db	2045	TGACTTGAATGCTCTCGGACATCGCTGGACGCGCAGGCCCTTCACGCTGGCAACAGAT	2104
QY	2733	GGTGCACGTGTGCTGGCAGCCGCAACAACAGCGTCTCCATGGTGAATTACAGATTCGAGT	2792
Db	2105	GGTGCACGTGTGCTGGCARTCGCAGACCAAGTGTCTCCATGGTGAATTACAGATTTGCTGT	2164
QY	2793	GGAGAAATCAGTGTCTGGAAAACTGCTGAGGAAATTTCAAAAAACGCAACGGGTGGCAGAA	2852
Db	2165	GGAGAAATCAGTCTCTGGGAACTTGTGAGAAAGTTCAAGAAACGCAATGGCTGGCAGAA	2224
QY	2853	GCTGTGGGTGTGTTTCAAACTTCTGCTGTCTTCTTCAAAATCACACAGACAATCA	2912
Db	2225	GCTGTGGGTGTATTCAACAACTTCTGCTGTCTTCTTCAAAATCAACACAGACAATCA	2284
QY	2913	TCCCTTTGCCAGCTGCTCTGTCTGGCTACTCGCTCAACTCCCTGTGATCCGAGAA	2972
Db	2285	CCCCCTTGGCCAGCTGCTCTGTGGGCTACTCCCTCAACAATCCCTCTGTGATCGAGAA	2344
QY	2973	CATCCAGAAAGATACGTGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGG	3032
Db	2345	CATCCAGAAAGATACGTGTGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGG	2404
QY	3033	GGAAAGCAGATACGTTTCGAAAGGTGATGGAAGTATCCGACAGTGCACACACTGTGC	3092
Db	2405	TGAAAGCAGATACATTTGAAAGGTGATGGAAGTATTCGAAAGTGCACACACTGCAC	2464
QY	3093	CTCGCGACCCGACGTGTGGAGCCCAAGAGTCTCTTGTGTGA	3134
Db	2465	CTCTCGGGCCCACTTGAAGTTCACAAAGATTCACCTGTGA	2506

LOCUS	COJ24059	2937 bp	DNA	linear	PAT 03-FEB-2004
DEFINITION	Sequence 9993 from Patent WO02068579.				
ACCESSION	COJ24059				
VERSION	COJ24059.1	GI:42284916			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;				
TITLE	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.				
JOURNAL	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.				
FEATURES	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof				
source	Patent: WO 02068579-A 9993 06-SEP-2002; PE Corporation (NY) (US)				
ORIGIN	Location/Qualifiers				
	1..2937				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
Query Match	59.1%; Score 1851.8; DB 2; Length 2937;				
Best Local Similarity	79.9%; Pred. No. 0;				
Matches 2513; Conservative	0; Mismatches 77; Indels 555; Gaps 6				
QY	172 CAAAGAGCTCCCTGGGAAGAGTGTCTGTGATCAGTTTGCACACACCTCAACCTCGGTGAA	231			
DB	1 CAAAGAGCTCTCGGGAAGAGTGTCTGTGATCAGTTTGCACACACCTCAACCTCGGTGAA	60			
QY	232 GGTGACTATTTTGGCGCTCGAGTTTCTGTATCAGAAAAAGATCAGCGTGTGCTGATCTC	291			
DB	61 GGTGACTATTTTGGCGCTCGAGTTTCTGTATCAGAAAAAGATCAGCGTGTGCTGATCTC	105			
QY	292 CTAAACCCATTGTGAAACAGATTGAAAGGCCAAAGCACGTTGTGTTAATTGTGTGTG	351			
DB	106 -----GCCAGGTCCCGGTG	120			
QY	352 AAATTCCTTCCGCGTCGACACACAACTCCAGAAAGACTCAGAGGTACTGTTCGCG	411			
DB	121 ACCTTCGCGCTCTTGTGCTGACGCCACGCTCTGTGATCTTCGACAGTACTGTTCGCG	180			
QY	412 CTGACAGTGAAGCAGAGCTTGCTCAAGGACGTTGACGTGTAAATGACACAGCGCAGCT	471			
DB	181 CTGACAGTGAAGCAGAGCTTGCTCAAGGACGTTGACGTGTAAATGACACAGCGCAGCT	240			
QY	472 CTCTTGATTTACACATTGTGCAATCTGAGATTTGGGATTTTGAATGAAGCCTTGTGACGA	531			
DB	241 CTCTTGATTTACACATTGTGCAATCTGAGATTTGGGATTTTGAATGAAGCCTTGTGACGA	300			
QY	532 GAGCACTTAGCAAAAAATTAATATCATCTCAGCAAGACGCACTAGAGGACAAATCGTG	591			
DB	301 GAGCACTTAGCAAAAAATTAATATCATCTCAGCAAGACGCACTAGAGGACAAATCGTG	360			
QY	592 GAATTTCAACATTAACAATTGACAAACACACAGCAGATCAAGTTTCACTCTCTAGAG	651			
DB	361 GAATTTCAACATTAACAATTGACAAACACACAGCAGATCAAGTTTCACTCTCTAGAG	420			
QY	652 ATTGCCGCTGGCTAGAGATGTAAGAAATCCGGTTGACCCGGCCAGAGGACAGGAAAGGC	711			
DB	421 ATTGCCGCTGGCTAGAGATGTAAGAAATCCGGTTGACCCGGCCAGAGGACAGGAAAGGC	480			
QY	712 ACGAAGATCAATCTGCGCGTTGCCAACAAGGAAATCTAGTGTTCAGGGTTTCACTAAG	771			
DB	481 ACGAAGATCAATCTGCGCGTTGCCAACAAGGAAATCTAGTGTTCAGGGTTTCACTAAG	540			
QY	772 ATCAATGCTTCAACTGGGCCAAGGTGGGAAGCTTAGAGCTTCAAGAGGAAGCGCTTTCTC	831			
DB	541 ATCAATGCTTCAACTGGGCCAAGGTGGGAAGCTTAGAGCTTCAAGAGGAAGCGCTTTCTC	600			

832 ATCAAGCTCCGGCCAGATGCCAATAGTGGTACCAAGATACCTTGAATTCCTGATGGCC 891
| | | | |
Db 601 ATCAAGCTCCGGCCAGATGCCAATAGTGGTACCAAGATACCTTGAATTCCTGATGGCC 660
| | | | |
Qy 892 AGTCGGGATTTTTCGAAGTCTCTTGGAAAATTTGTGTTGAACATATGCTTCTTTAGA 951
| | | | |
Db 661 AGTCGGGATTTTTCGAAGTCTCTTGGAAAATTTGTGTTGAACATATGCTTCTTTAGA 720
| | | | |
Qy 952 CTTTTTGAAGAGCCCAACCAAGCCCAAGCCCGTCTCTTTAGCCGGGGGTATCATTTT 1011
| | | | |
Db 721 CTTTTTGAAGAGCCCAACCAAGCCCAAGCCCGTCTCTTTAGCCGGGGGTATCATTTT 780
| | | | |
Qy 1012 CGATTTCAGTGTTCGACTCAGAAAGCAGGTTCTCGACTATGTTAAAGAAAGAGACATAAG 1071
| | | | |
Db 781 CGATTTCAGTGTTCGACTCAGAAAGCAGGTTCTCGACTATGTTAAAGAAAGAGACATAAG 840
| | | | |
Qy 1072 AAGGTGCAGTTTGAAGAAAGCAAGCAAGATTCAATTCATCCGAGCCTTGTTCACAG 1131
| | | | |
Db 841 AAGGTGCAGTTTGAAGAAAGCAAGCAAGATTCAATTCATCCGAGCCTTGTTCACAG 900
| | | | |
Qy 1132 CCTACAGAACTGAATTCCGAAAGTGTGAGCAGTCTCAGAGAGACCAAGCCTTACATTT 1191
| | | | |
Db 901 CCTACAGAACTGAATTCCGAAAGTGTGAGCAGTCTCAGAGAGACCAAGCCTTACATTT 960
| | | | |
Qy 1192 GGAGAAAGTGCAGAAATCTCAAGGGGGCCAGAGCTGCAGAGAAAGAAACCGAAGGTT 1251
| | | | |
Db 961 GGAGAAAGTGCAGAAATCTCAAGGGGGCCAGAGCTGCAGAGAAAGAAACCGAAGGTT 1020
| | | | |
Qy 1252 TCCGCGGGGAGCCGGGGGTTCGACCCGAGCCTTGCAGAGAAAGCCCGGGGTAAAC 1311
| | | | |
Db 1021 TCCGCGGGGAGCCGGGGGTTCGACCCGAGCCTTGCAGAGAAAGCCCGGGGTAAAC 1080
| | | | |
Qy 1312 AAGCAGCCGAGCCGAGCCGCTCGAGCCCAAGAGAAAGAGAGAGGTCGTTAAGAT 1371
| | | | |
Db 1081 AAGCAGCCGAGCCGAGCCGCTCGAGCCCAAGAGAAAGAGAGAGGTCGTTAAGAT 1140
| | | | |
Qy 1372 AAGCAGCCGAGCCGAGCCGCTCGAGCCCAAGAGAAAGAGAGGTCGTTAAGAT 1413
| | | | |
Db 1141 AAGCAGCCGAGCCGAGCCGCTCGAGCCCAAGAGAAAGAGAGGTCGTTAAGAT 1200
| | | | |
Qy 1414 ----- 1413
| | | | |
Db 1201 GCCAGAGGAGCAATTCATTCCTTTCAATTGACTGAGTGAAGTGAACGAATAC 1260
| | | | |
Qy 1414 ----- 1413
| | | | |
Db 1261 GACCTTCTTAGAAGAAACAAGCTCCGATCCCGATCCCGATCCCGATCCCGATCCCGATCC 1320
| | | | |
Qy 1414 ----- 1413
| | | | |
Db 1321 AATGACAAATTATGAGGGCTCCCTGACTGGGAGTCTCACCTTTCCGAGCTGTGTGAAC 1380
| | | | |
Qy 1459 TCGCAGGGGGAGTGGCCCCCTGCGCAAGTGAATTTCTTCCCACTTGAGCCCGACAC 1518
| | | | |
Db 1381 TCGCAGGGGGAGTGGCCCCCTGCGCAAGTGAATTTCTTCCCACTTGAGCCCGACAC 1440
| | | | |
Qy 1519 AAGCAGGCTCTCCCTTGAATGAGCCGCTGCTGAATGACCAAGGCTGCCCGCGAGAC 1578
| | | | |
Db 1441 AAGCAGGCTCTCCCTTGAATGAGCCGCTGCTGAATGACCAAGGCTGCCCGCGAGAC 1500
| | | | |
Qy 1579 GATGAGATGAGGGCCGAGAGAGAGATTTCCCACTGATTAAGCGTACTTCAATGCTAAG 1638
| | | | |
Db 1501 GATGAGATGAGGGCCGAGAGAGAGATTTCCCACTGATTAAGCGTACTTCAATGCTAAG 1560
| | | | |
Qy 1639 GAAAGTCTTACCAACCGAGGAAATATCTGAAAGATCTGAAAGTATCTGCTGTGTTT 1698
| | | | |
Db 1561 GAAAGTCTTACCAACCGAGGAAATATCTGAAAGATCTGAAAGTATCTGCTGTGTTT 1620
| | | | |
Qy 1699 CAGAGCAGTGAAGCAAG 1758
| | | | |
Db 1621 CAGAGCAGTGAAGCAAG 1680
| | | | |
Qy 1759 AATTTTGAACCTTTGCAAAATTTCACTAATTTTCTCAAGAAATTTGAGCAACGACTT 1818
| | | | |

1681 AATTTTGAACCTTTGCAAAATTTCACTAATTTTCTCAAGAAATTTGAGCAACGACTT 1740
| | | | |
Qy 1819 GCCCTGTGGAGAGGCCGCTCAAAATGCCAAATCAAGATTTACCAAGAAATCCGCGATGTC 1878
| | | | |
Db 1741 GCCCTGTGGATGAGCTGACA----- 1761
| | | | |
Qy 1879 ATGCTGAAGAACATTCAGGGCATGAAGCACTGCGGCTCACCTGTGGAAGCAGACGAG 1938
| | | | |
Db 1762 ----- 1761
| | | | |
Qy 1939 GCTTTGAGAGCCCTGAGAGATGAATCAAGAGCTCCCGCGGCTGAGAACTTTGCGACA 1998
| | | | |
Db 1762 ----- GCTGTGAGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1786
| | | | |
Qy 1999 GACTTTGAGCTGAGAGAGTGTGTTAAGTACCGCTCAACACTTCTCTGTGGGCACTG 2058
| | | | |
Db 1787 CTTTCCGCTGCAACCTGTATGTCTTTC----- 1817
| | | | |
Qy 2059 CACCGGCTCATGCACTAAGCAGGTCCTGAGCGGCTGTGCAAAACACCAAGCCGCGAGC 2118
| | | | |
Db 1818 ----- 1817
| | | | |
Qy 2119 CAGCCGACTTCAGGAGACTGCGAGCCGCTTTGGCAGAGATCAAGAGATGCTGCAAG 2178
| | | | |
Db 1818 -----CTTCACAGCCGCTTTGGCAGAGATCAAGAGATGCTGCGACAG 1860
| | | | |
Qy 2179 CTCCAAGTATGATGATCAAGATGAGAAATTTCCAGAAAGCTGCAAGAACTCAAGAAAGAT 2238
| | | | |
Db 1861 CTCCAAGTATGATGATCAAGATGAGAAATTTCCAGAAAGCTGCAAGAACTCAAGAAAGAT 1920
| | | | |
Qy 2239 TTGATTTGATGATCAATCTTGTGTTCCGGGAGAGGAGTTCAATCGTCTGGGAGCCCTC 2298
| | | | |
Db 1921 TTGATTTGATGATCAATCTTGTGTTCCGGGAGAGGAGTTCAATCGTCTGGGAGCCCTC 1980
| | | | |
Qy 2299 AAGCAAGCTCTCGGGAGAGGGGCTCCAGAGCGCATGTTCTTCTCTG----- 2343
| | | | |
Db 1981 AAGCAAGCTCTCGGGAGAGGGGCTCCAGAGCGCATGTTCTTCTGTGATGAGAGAGAGC 2040
| | | | |
Qy 2344 ----- 2343
| | | | |
Db 2041 GGCTTGTCTCAAGAAATTTGTGTCACTGGGCAACAAAGCTCCCAAGCCGAGCTTCG 2100
| | | | |
Qy 2344 -----TTCAAGAGCTCTGCTATACAGAGCCGGGGCTGACGGCTTCAATCAG 2394
| | | | |
Db 2101 GCCACCAATTCAAGAGAGTCTGCTATACAGAGCCGGGGCTGACGGCTTCAATCAG 2160
| | | | |
Qy 2395 TTTTAAAGTCCACGGGAGCTCCGCTCTATGAGCATGAGATTGAGAGAGAGAGAGAGAG 2454
| | | | |
Db 2161 TTTTAAAGTCCACGGGAGCTCCGCTCTATGAGCATGAGATTGAGAGAGAGAGAGAGAG 2220
| | | | |
Qy 2455 TGGGGGGTGGCCCACTGCTGACCTCCGGGGCCAGCGGAGTCCATCATCGTGGCCGCTC 2514
| | | | |
Db 2221 TGGGGGGTGGCCCACTGCTGACCTCCGGGGCCAGCGGAGTCCATCATCGTGGCCGCTC 2280
| | | | |
Qy 2515 AGTTCTCGTCCGAGATGAGAGAGTGGTGAAGCATTCAGATGAGCAATGACCTGGCG 2574
| | | | |
Db 2281 AGTTCTCGTCCGAGATGAGAGAGTGGTGAAGCATTCAGATGAGCAATGACCTGGCG 2340
| | | | |
Qy 2575 GAGAAAGAGAGAGAGCCCGCCCTGAGTTCCTGAGCAGAGCCCGCTGAGCAACAGTCC 2634
| | | | |
Db 2341 GAGAAAGAGAGAGAGCCCGCCCTGAGTTCCTGAGCAGAGCCCGCTGAGCAACAGTCC 2400
| | | | |
Qy 2635 CCGATGAAGCCACCGCGGCTGACAGGAGTCAAGAGATGACCTGAGCGCCTCGCGACA 2694
| | | | |
Db 2401 CCGATGAAGCCACCGCGGCTGACAGGAGTCAAGAGATGACCTGAGCGCCTCGCGACA 2460
| | | | |
Qy 2695 TCGCTGAGAGCCGAGAGCCCGGACCGAGCAACAAATGTGTCAAGTGTGTGAGCAGCGC 2754
| | | | |
Db 2461 TCGCTGAGAGCCGAGAGCCCGGACCGAGCAACAAATGTGTCAAGTGTGTGAGCAGCGC 2520
| | | | |
Qy 2755 AAGACCAAGCTTCAATGTGAGCTTCAGCA-----TCCAGTGGAG 2796
| | | | |

Db		2821	AACACGAGCGTCTCCATGTGTGGAGACTTCAGCAATTGCAGTAGGAGGCTCTTGCCTTTAG	2598
Oy		2797	AATCAGTTGTCTGGAAACCTGTGTGAAGAAATTCAAACAAGCACMGSGGTGGCAGAAGCTG	2856
Db		2581	AATCAGTTGTCTGGAAACCTGTGTGAAGAAATTCAAACAAGCACMGSGGTGGCAGAAGCTG	2640
Oy		2857	TGGGTGGTGTTCACAAAATTCTGCGCTGTCTTTCTTAAGATAACACACAGAACATCATGCC	2916
Db		2641	TGGGTGGTGTTCACAAAATTCTGCGCTGTCTTTCTTAAGATAACACACAGAACATCATGCC	2700
Oy		2917	CTTGCACGACCTGCGCTCTGCTGCGGCTACTCGGCTCACACATCCCCTCTGAATCCGAGAACATC	2976
Db		2701	CTTGCACGACCTGCGCTCTGCTGCGGCTACTCGGCTCACACATCCCCTCTGAATCCGAGAACATC	2766
Oy		2977	CAGAAAGACTACGTGTTCAGCTGCACCTTCAAGTCCACGTCCTACTAATTAGAGGCGGAA	3036
Db		2761	CAGAAAGACTACGTGTTCAGCTGCACCTTCAAGTCCACGTCCTACTAATTAGAGGCGGAA	2820
Oy		3037	AGCAGTACACAGCTTCGAAAGGTGGA	3061
Db		2821	AGCAGTACACAGCTTCGAAAGCATGA	2845
RESULT_9				
BC004009				
LOCUS				
DEFINITION		BC004009	2632 bp mRNA linear ROD 21-OCT-2003	
VERSION		Mus musculus FERM, RhosGEF (Arfge) and pleckstrin domain protein 1		
KEYWORDS		(Chondrocyte-derived), mRNA (CDNA clone IMAGE:3493093), partial cde.		
ACCESSION		BC004009		
ORGANISM		GI:13278387		
SOURCE		Mus musculus (house mouse)		
REFERENCE		Mus musculus		
AUTHORS		Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Mus.		
		1 (bases 1 to 2632)		
		Strauberg,R.L., Feltingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,H., Shenmen,C.M., Schuler,G.D., Altschuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhac,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.N., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Uedin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McKean,P.J., McKernan,K.U., Malek,U.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.U., Huliyil,S.W., Vallajon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,R., Whiting,M., Madan,A., Young,A.C., Shechenko,Y., Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Gilmwood,J., Schmutz,J., Myers,R.M., Butcherfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E., Schermer,A., Schein,J.E., Jones,S.J. and Marra,M.A.		
		Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
TITLE		Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
JOURNAL		12477932		
PUBMED		2 (bases 1 to 2632)		
REFERENCE		Strauberg,R.		
AUTHORS		Direct Submission		
JOURNAL		Submitted (28-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK		NIH-MGC Project URL: http://mgc.ncl.nih.gov		
COMMENT		Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>
Series: IRAK file: 8 Row: P Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

```

source
1. 2632
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech 11"
/db_xref="taxon:10090"
/clone_xref="IMAGE:3493093"
/tissue_type="Mammary tumor metastasized to lung, Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMV."
/clone_1ib="NCI_CGAP_L429"
/lab_host="DH10B"
/lab_name="Vector: pCMV-SPORT6"
<1. 2632
/misc_feature
/feature="Parp1"
/feature="synonym: CdeP"
/db_xref="geneid:223254"
/db_xref="MGI:2446173"
<1. 1263
CDS
/feature="Parp1"
/codon_start=1
/product="Parp1 protein"
/protein_id="AAH04009.1"
/db_xref="GI:13278388"
/db_xref="geneid:223254"
/db_xref="MGI:2446173"
/translation="VMLKNTQGMKHLAAHLMKHSEALEALETSIKGSRRLHFCRDE
LOKQVLEPLNTFLPLRLRLMRYGVHLERLCKNHPRHADPPCRRAALAEITENVQL
HETMKMNFQKLEHLKDLIGDNLVLPKEREFTGLSDLSKSLQGLQAFPLFNDV
LLYTSRGITLSEINOKRVGQPLPYMTLIESEEEGVHCVTLTGQRSIIIVASSRE
MEKMWEDLOMAIDLAESKNGPTPELLASSPPDNKSPDEATPAQOESHDIDLASRTSL
ROAPRGSDTHV/CWHSRTSVWDFSLAVENOLSGVLLKRFKSNQMOGLVWFTFPL
CLPFKSGDHSPLASLPILGYSLLTPESENIHKOVVFLKHFPSHYVIRASSEYTF
ERMEVITRSATSSASRAHITLSKESHLY"
misc_feature
55. 309
/feature="Parp1"
/feature="RhoGEF; Region: Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPases"
/db_xref="CDD:smat00325"
448. 659
/misc_feature
/feature="Parp1"
/feature="PH; Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syncrophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"
/db_xref="CDD:smat00233"
928. 1218
misc_feature
/feature="Parp1"
/feature="PH; Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various

```


proteins. PH domains have been found to possess inserted domains (such as in p13C gamma, syntrophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol 1phid8"

/db_xref="CD:smart00233"

ORIGIN

Query Match 33.3%; Score 1043; DB 6; Length 2632;

Best Local Similarity 89.3%; Pred. No. 2.6e-282;

Matches 1124; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

```

QY 1876 GTCATGCTGAAGAACTTTAGGGCATGAAGCACTTGGCGGCTCAGCTGTGGAAAGCAGC 1935
DB 1 GTCATGCTGAAGAACTTTAGGGCATGAAGCACTTGGCGGCTCAGCTGTGGAAAGCAGT 60
QY 1936 GAGGCTTGGAGGCTTGGAGATGGATCAAGAGCTCCGGCGGCTGGAGAACTTCTG 1995
DB 61 GAGGCTTGGAGGCTTGGAGATGGATCAAGAGCTCCGGCGGCTGGAGAACTTCTTGC 120
QY 1996 AGAGACTTTAGCTGAGAGAGGTGTCTTACCTACCGCTCAACACCTTCTCTGGGCGCA 2055
DB 121 CGAGACTTGAAGTGGAGAGAGGTGTCTTACCTACCGCTCAACACCTTCTCTGGGCGCG 180
QY 2056 CTGCACCGGCTCATGATCACTCAAGAGAGTCTTGGAGAGGCTGTGGAAACCAACCGCGG 2115
DB 181 CTGCACCGGCTCATGATCACTCAAGAGAGTCTTGGAGAGGCTGTGGAAACCAACCGCA 240
QY 2116 AGCCAGCGGACTTTCAGGAGCTGCGAGCGCTTTGGGAGATCAAGAGATGGTGGCA 2175
DB 241 AACCAAGCGGACTTTCAGGAGCTGCGAGCGCTTTGGGAGATCAAGAGATGGTGGCG 300
QY 2176 CAGCTTCACGATGATGATCAAGATGAGAAATTTCCAGAGCTGCACGAACTCAAGAA 2235
DB 301 CAGCTTCACGATGATGATCAAGATGAGAAATTTCCAGAGCTGCATGATGATCAAGAA 360
QY 2236 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2295
DB 361 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
QY 2296 CTACAGAGCTCTCGGGGAGAGGGGCTCAAGAGCGCATTTCTTCTGTTCAACGAGTC 2355
DB 421 CTACAGAGCTCTCGGGGAGAGGGGCTCAAGAGCGCATTTCTTCTGTTCAACGAGTC 480
QY 2356 CTGCTATACAGAGCGGGGGCTGACGGCTTCAATCAATTAAAGTCCACGGGAGCTC 2415
DB 481 TTGCTGTATACAGCGGGGGAGCTGACAGCATCTAATCAATTAAAGTCCACGGGAGCTC 540
QY 2416 CGGCTCTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2475
DB 541 CAGCTCTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 2476 ACCCTCCGGGGCAGCGGAGTCTCATCATCGTGGCGGCACTTCTCGATCCGAGATGAG 2535
DB 601 ACCCTCCGGGGCAGCGGAGTCTCATCATCGTGGCGGCACTTCTCGATCCGAGATGAG 660
QY 2536 AAGTGGGTTGAGAGATCCAGATGGCATTTGACTTGGCGGAGAGAGAGAGAGAGAGAG 2595
DB 661 AAGTGGGTTGAGAGATCCAGATGGCATTTGACTTGGCGGAGAGAGAGAGAGAGAGAG 720
QY 2596 CCTGAGTTCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2655
DB 721 CCTGAGTTCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 2656 GACCAAGAGTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2715
DB 781 GACCAAGAGTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 2716 CACCGGAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2775

```

```

DB 841 CACCGGCAACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 2776 GACTTCAGCATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2835
DB 901 GACTTCAGCATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 2836 AGCAACGGGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2895
DB 961 AGCAACGGGTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 2896 TCACACGAGATCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 2955
DB 1021 TCACACGAGATCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1080
QY 2956 CCTCTGAGTCTGAGAGATCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 3015
DB 1081 CCTCTGAGTCTGAGAGATCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1140
QY 3016 GTCTACTCTTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3075
DB 1141 GTCTACTCTTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 3076 AGTCCACAGAGTCTGCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3134
DB 1201 AGTCCACAGAGTCTGCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1259

```

RESULT 10

LOCUS CQ725044

DEFINITION Sequence 10978 from Patent WO02068579.

ACCESSION CQ725044

VERSION CQ725044.1 GI:42285901

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

Patent: WO 02068579-A 10978 06-SEP-2002;

PE Corporation (NY) (US)

FEATURES

source

1. .3997

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 33.1%; Score 1039; DB 2; Length 3997;

Best Local Similarity 60.3%; Pred. No. 3.9e-281;

Matches 1865; Conservative 0; Mismatches 1185; Indels 45; Gaps 7;

```

QY 21 GCCGACCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 80
DB 150 GCGAGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 209
QY 81 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 140
DB 210 TGGGAGAGCTCTTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 269
QY 141 GAGTACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 200
DB 270 GAGACAAACCAATGAGAAATTTGACATTTGACATTTGACATTTGACATTTGACAT 329
QY 201 TGAGATTGCAACCACTCACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 260
DB 330 ACAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 389

```

Qy 261 TCACAAAAGATCAGCGGTGGCTGATCTCCTAAACCCATTTGTGAACGATTAAG 320
Db 390 TACTAGTCTTACTGGAATTTGGCTTGAACCTATGAACCATCTTTAGGCAATACAG 449
Qy 321 GCCAAGACAGTTGTTGTAAGTTGTGAAATTCCTTTCCGCTGACACACAACACT 380
Db 450 GCCAAGAAATGTGGTCTCGCTAGCTGAATTTTTCACCTGATCCTGGTACGCT 509
Qy 381 CCAGAAGAACTCACAAGGTAACCTGTTCCGGCTGCAAGTGAAGAGAGACTTGGCTCAAG 440
Db 510 ACAAGAAATATATACAGATATCTTGTTCCTTGCACTTAAGAGAACTGCTGAGA 569
Qy 441 CAGGTGACGTATATGACACAGCGAGCTCTTGATTTCAACATGTCGAATGTA 500
Db 570 GCGTTGACCTGTGCTGACACACAGCGGCTTCTCAAGTCCCATCTCTGCACTGGA 629
Qy 501 GATTGGGATTTTGTATGAGCCTTGAACAGAGACACTTGAACAAAATTAATACATACC 560
Db 630 AATAGGATTTACGATGAACGCTGACCGAGAGCACTCAAGTGAACGATATTTGCC 689
Qy 561 TCAGCAAGACGCACTAGAGACAAATCGTGAATTTCAACATTAACCACTTGGACAAAC 620
Db 690 TGGCCAGACGACACTCCCTTGAAGAAATTAATTCATCAGAAAGCACTGGGCCAGAC 749
Qy 621 ACCAGCAGAATCAGATTTCCAGCTCTAGAGATTTGCCGCTGAGATGATGGAAT 680
Db 750 ACCTGCTGAGTCCGATTTCCAGGTGCTCGAAATTTGCTCGAAATTTGAATTAACGCAAT 809
Qy 681 CCGGTTGCAACCCGCGCAAGAGAGAGAGAGAGAGAGAGATCATCTGGCCCTGCAACAC 740
Db 810 CAGATTTCAATGCTTCTGACAGGGAAGAGACCAAGATTCATGCGCACTTTCCCAT 869
Qy 741 GGGAAATCTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCACTGGGCCAAGTGG 800
Db 870 GGGTGAATCTGTTTCCAGGCAACCAAAATCAACATTTCACTGCTCCAGGCTCG 929
Qy 801 GAAGCTGAGCTTCAAG 860
Db 930 TAAACTAGCTTCAAG 989
Qy 861 GTAACAGAGATCCTTGAATTCCTGATGCGAGTCCGAGATTTCTGCAAGTCTTCTGAA 920
Db 990 TTAACAGAGACATTAAGATTTTGTGGTAGAGATGAATGAATGAATCTTCTGAA 1049
Qy 921 AATCTGTGTGAACATCATGCTTCTTTTGAACCTTTTGAAGAGCCCAACCAAGCCCA 980
Db 1050 GATTGTGTGAGATCAACCTTTTGTAGACTTTTGAACCAACTTAAGCCAAACCAAA 1109
Qy 981 GCCGCTCTCTTGAAGCGGGGGTCACTATTCGAGTTCAGTGTGCGACTGAGAGAGGT 1040
Db 1110 AGCCGCTCTTGAAGCGGGGGTCTCTCTTCAATACAGTGAAGAGAGAGAGAGAGAG 1169
Qy 1041 TCTGCACTATGTTAAAG 1100
Db 1170 AGTAGATTTATTCAG 1229
Qy 1101 GATTCTTCTATCCGAGAGCTTCTTCAACGCTTCAAGAACTGAATTCGAGAGTCTGGA 1160
Db 1230 GACCCACAGGTCC-----GTTGAGGCTCTGACGTGACAGACTTACCAAA 1271
Qy 1161 GCAGTCTCAG 1220
Db 1272 ACAGAGACTCTATTTCCCGAGAGAGATTAAGAGACTCTGCTCCCATCTTCAAGAGATGC 1331
Qy 1221 GAGCTGCGCGAG 1280
Db 1332 CTTTTCAGGCTCTCTCCCTCACTCTGAGTCCCTCTGAGCTGCAAGATTTAAGAGAG 1391
Qy 1281 CCTTGCGCGAG 1340
Db 1392 CAGCAGACTCTCTCAAGATCCAGAGTTCCTAGTCAAGAGTCCAGCTGCAAGAGAGG 1451
Qy 1341 CACGAGAGAGAGAGAGAGAG-----TCGTTAAGATTAAGAGAGAGAGAGAGAGAG 1397

Db 1452 CAGTGAAGAGAGTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1511
Qy 1398 CCGCAG 1457
Db 1512 ACTCAG 1571
Qy 1458 CTGCGAG 1517
Db 1572 GAGCCCCCTGAGTCTGAGAGCTCTGATTTCAAGTGCCTT-----TGGGCCACCTGA 1622
Qy 1518 CAAGCAG 1577
Db 1623 ACAG 1682
Qy 1578 CGATGAG 1637
Db 1683 CTGCGAG-----GAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1739
Qy 1638 GGAAGTGTCTACACCGAG 1697
Db 1740 AGAGATTTCTGCTACAG 1799
Qy 1698 TCAGAGCAG 1757
Db 1800 CCGCAG 1859
Qy 1758 GAAATTTGAACCTTTGACAAATTTATTAATTTCTCAAGAGAGAGAGAGAGAGAG 1817
Db 1860 CAACATCGATCCCATCTATAGAGTTTCCAGAGAGCTTCTGCGGAGAGAGAGAGAGAG 1919
Qy 1818 TGGCTGTGGAG 1874
Db 1920 GGCACCTGAG 1979
Qy 1875 TGTCTGTGAG 1934
Db 1980 CATCTGCTCAG 2039
Qy 1935 CGAGGCTTGAAG 1994
Db 2040 CGAGGCTCTTAACAG 2099
Qy 1995 CAGAGACTTGAAG 2054
Db 2100 CAG 2159
Qy 2055 ACTGCAAG 2114
Db 2160 CATCAG 2219
Qy 2115 GAGCAG 2174
Db 2220 CCGGACAG 2279
Qy 2175 ACAGCTCAG 2234
Db 2280 CACACTACAG 2339
Qy 2235 AGATTTGATTTGAGATGAG 2294
Db 2340 GAG 2399
Qy 2295 CCTGAG 2354
Db 2400 CTTTCAAG 2459
Qy 2355 CTGCTATACAG 2414
Db 2460 GTTGTCTGACAG 2519
Qy 2415 CCGGCTCTATGAG 2474

Db 2520 TCCCTCAAGGAGATCTGTGGTGAAGAAAGATGAACGATGTCTGTTCCACACTGTTT 2579
QY 2475 GACCTCCCGGGGCGACGGGCAATCCATTCGTGGCCGCAATTCCTCGGTCCGAGATGGA 2534
Db 2580 CACCAATCTAAGGGCTCAAGAAAACAATCGTGGGCGAGCAGACACTCGGCTGGAAGAA 2639
QY 2535 GAAGTGGTGGAGGACATCCAGATGSCATTGACCTGGCGGAGAGAGAGAGAGAGAGAGAG 2594
Db 2640 GAAGTGGATGCTGGACCTGAATCTCCGCGATTCAGAGAGCCAGAGAGAGAGAGAGAG 2699
QY 2595 CCTGAGTCTTGGCCAGCA-----GCCCTCCCTGCAACAAGTCCCTGATGAGAGCCAC 2648
Db 2700 CCTGCACTGCAAGGCGCGCACTGTGTGCACTGCTCCCGCCAGATCCCGCAAGAGGTATC 2759
QY 2649 CCGCGGTGACAGAGATCAGAGATGACCTGAGCGCTGCGGCACTGGCTGGAAGGCGCA 2708
Db 2760 TCTG---GAGCAGAGATCAGAGATGATGCTCGGGGTGTCCGAGCTCCCTGGAAGGCGCA 2816
QY 2709 GCGCCCGCAGCGCGGAGACACAATGCTGACGCTGTGCTGSCACCGCAGACAGACAGCTCTC 2768
Db 2817 TGGCCAGCAGCGGGCCACACCAATGACAGCTGTGCTGTACCGGAAACAGAGCTGTG 2876
QY 2769 CATGTGATCTTACAGATCGAGTGAAGATCATGTTGTTGAAACCTGCTGAGAAATT 2828
Db 2877 CAGGCGAGACCAAGTGCAGCTGTGAGAAACAGCTTTCAGAGATCTGCTAAGAAAGTT 2936
QY 2829 CAAAAACAGCAACGGGTGGCAAGAGCTGTGGGTGTGTTCAACAACTTTCGCTGTTCTT 2888
Db 2937 CAAAAACATCATGCTGTGAGAGAGCTGTGGGTGTGTTTACCAACTTGTGTTGTTCTT 2996
QY 2889 CTACAATACACACAGAGACATATATCCCTTGGCAGCTGCTGTGCTGAGCTACTGCT 2948
Db 2997 CTACAAAATCATGCTGTGAGAGAGCTGTGGGTGTGTTTACCAACTTGTGTTGTTCTT 2996
QY 2949 CACCATCCCTCTTGAATCCGAGAACATCCAGAAAGATCTACGTGTTCAACCTGACCTTCA 3008
Db 3057 GAGCATCCCGCAGGAGGCGAGATGSCATGACAAAGATCTATGTTTCAACCTCAGTTTCAA 3116
QY 3009 GTCCCAAGCTCTACTTTCAGGCGGAGAGAGAGATGACAGTTGCAAAAGGTGATGGAAT 3068
Db 3117 ATCCCAAGCTCTTCTTCCGGGCTGAGAGCAAGTACAAATTTGAAAGGTGATGAGGT 3176
QY 3069 GATCCGAGTGGCCACAGCTGTGCTTGGGACCC 3103
Db 3177 GATCCAGGGGGCCAGCAGCTCAGCCGGGAGGGGCC 3211

RESULT 11

LOCUS CO895396 3997 bp DNA linear PAT 05-NOV-2004
DEFINITION Sequence 16 from Patent EP1471153.
ACCESSION CO895396
VERSION CO895396.1 GI:55467577

KEYWORDS

ORGANISM Homo sapiens (human)
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE

1. Certi, U., Foser, S. and Weyer, K.
Transcriptional activity assay
Patent: EP 1471153-A 16 27-OCT-2004;
F. HOFMANN-LA ROCHE AG (CH)
Location/Qualifiers

FEATURES
source 1..3997
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 33.1%; Score 1037.4; DB 2; Length 3997;
Best Local Similarity 60.2%; Pred. No. 1.1e-280;

Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

QY 21 GCCGACCCAGGATACGACTGCGGGGCGCCGGAAAAATTCCGGGATCAGTACTTGAACG 80
Db 150 GCGACTGAGGATGCGCTTGGGTGCGAGCCCTGTGGAGTTAGACCTTGAAGCC 209
QY 81 TGGACAGAGCGCCCGCCCAACCTTCAAGAAATCTGTGTTCATCAAAATCAGATGCT 140
Db 210 TGGGAGAGCTCTTTGCCAGAAATGCAAGAGAGACCTGCACCTCAGATGAAGCTGCT 269
QY 141 GAGTACACCCAGAGAGGATTTGAAGTTTCACAAAGAGCTCTGGGAAAGTCTGTGGA 200
Db 270 GAGCAACACATGGAATAATTTGACATTGAGCTTAAAGCGATGCGCATTTACTGAC 329
QY 201 TGCAGTTTGAACAACCACTCAACTGATGAAAGATGATATTTGGCTGAGTTTCTGGA 260
Db 330 ACAGTGTGAAGGCTTTTAACTGTGATGATGTACTACTTGGGATGAGTTCAAAA 389
QY 261 TCACAAAAAGATCACGCTGAGCTGAGATCTCTTAAACCCATTTGTGAACAGATTAGA 320
Db 390 TACTGAGTCTACTGAGATTTGGCTTGAACCTATGAAACCATCATTTAGCAAAATGAGA 449
QY 321 GCCAAGACAGTTGTTGTTAAATTGTGTGAATTTCTTTCGCTGACACACCACT 380
Db 450 GCCAAGAGATGTGTGCTTGTGCTGATGCTGTAAATTTTTCACCTGATCTGTGACGT 509
QY 381 CCAAGAAATCTCAAGATGACTGTGCGCTGAGGTAAGAGACAGACTTGTGCTCAAG 440
Db 510 ACMAAGAAATTAAGAAATCTTGTGCTTGTGCACTTAAAGAGACCTGTGGAAGA 569
QY 441 CAGTTGACGTGTAAATGACACGAGGAGCTCTTGTGATTTTCACATTTGTGCAATCTGA 500
Db 570 GGGTTTGAAGTGTGTGACACACAGCGGCCCTTCTCAGTCCATCTCTGCACTGGA 629
QY 501 GATTGGGATTTTGAATGAAAGCTTGTGAAGAGAGACCTTGAAGAAATTAATATAC 560
Db 630 AATAGAAATTAAGATGAAGAGCTGTGAGACCGAGAGACCTCAAGATGAAGATTTTGGC 689
QY 561 TCAGCAAGCGCATAGAGAGCAAAATGTGGAATTTACCACTAATCACTTGAACAA 620
Db 690 TGGCAGACGACTGCTTGAAGAAATCTAGAAATTCATCAAGAGCAGTGGGCGAGAC 749
QY 621 ACCAGCAATACAGATTTTCAGCTCTTGAAGATTTGCCGTGGCTAGAGATGATGAAT 680
Db 750 ACTGCTGATGAGATTTCAAGTGTCTGAATTTCTGAAAGTTGGAATGTACGGCAT 809
QY 681 CCGGTGACCCCGGCAAGAGACAGGAGAGCAAGAAATCAATCTGGCCGTTGCCAACAC 740
Db 810 CAGATTTACATGCGCTTCTGACAGGAGAAACCAAGATTCAACTGACAGTTCACAT 869
QY 741 GGGATTTCAAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGAGGCAAGGTGCG 800
Db 870 GGGTGTACTGTTGTTCCAGGGGACACCAAAATCAACCTTCAACTGTGTCCAAAGTCCG 929
QY 801 GAGCTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGGCCAGATGCCAATAGTGC 860
Db 930 TAAACTAAGCTTCAAGAGAAAGATTTCTTATCAAACTTCACTCAGAGGTTCAAGACC 989
QY 861 GTACAGAGATCTTGGAAATTCGTAGATGCGCAGTGGGATTTTTCAGAGTCTTCTGGA 920
Db 990 TTACAGAGACATTAAGAAATTTGTGGTGAAGATGAATTAAGAACTTCTGGA 1049
QY 921 AATCTGTGTAACATCATGCTTCTTATAGACTTTTGAAGGCCCAACCAAGCCCA 980
Db 1030 GATTTGTGTGAGATATCACACTTTTATGACTTTTGAACCACTTAAGCCAAAGCAAA 1109
QY 981 GCGGTCTCTTTAGCCGGGGGTCTATCTTCTGATCAAGTGTGAGTCTCAGAGCAGGT 1040
Db 1110 AGCGTCTTCTTCAAGCGGGGCTCTCTCAGATACAGTGAAGAACTCAGAAAACT 1169
QY 1041 TCTGACATATGTTAAAGAGAGACATTAAGAGGTGACGTTTGAAGAGAGACAGCA 1100
Db 1170 AGTAGATTATTTCAAGACAGTGAATGAAGAAATTCATATGAAGAGGACAGCA 1229

QY 1101 GATTCTTCTATCCGAGCCTTGCTTACAGCCTTACAGAACTGAAATTCGGAAGTGTGGA 1160
DB 1230 GACCCACAGTCC-----GTTTCAGCTCTGACATGCAACCTTACCAAA 1271
QY 1161 GCAGTCTCAGACAGACACCACTTACATTTGGAAGAGTGCAGAAATCTTCAGAGGGCCA 1220
DB 1272 ACAGACATCTCATTTCCCGAAGGGAATTGAGGACTCCCTGCTCCCACTTTCACAGAAATGC 1331
QY 1221 GAGCTCCGGCGAGAAAGAACCCGAAGTTTCCGCGGGAGCCGGGGTGCACCCGAG 1280
DB 1332 CTTTTACTGCTCTTCTCCCTCCACTCTGATCTCCCTGCGCTCCAGAGTTTAAAGACAG 1391
QY 1281 CCTCGGCGGAGAGAACCCCGCGGTAAACAAGACGGAGCGAGCCGCTTCGCGCC 1340
DB 1392 CAGCAGCTCCCTCAGATCCCAAGTTTCTTACGTCAAGATGTCAGCTGCAGAGAGGCG 1451
QY 1341 CACGGAAGAAAGAGAGAG---TCCGTTAAGATAGGACCCAGACAGATTAACCTGAGCC 1397
DB 1452 CAGTGAAGCAGTGGCTGAGAGCCCCCAGACACCATCGGCCAGCCCTCGGGGCCCCCGC 1511
QY 1398 CCGCAGCCAGACAGACAGCTCCCTGACTGGCAGTCTCACCTTTCGAGCTGTCTGAA 1457
DB 1512 ACTCCAGCTGTGTCAGAGCCTTTCAGAGAGTCTCAGCCTTCTCCCTCCAGCGGAA 1571
QY 1458 CTCGCAAGGGGGAGTGGCCCCCTTGCACAGTACCTTGTCTCCCAACTGAGCCCCGAC 1517
DB 1572 GAGCCCCCTGAGTCTAGAGCCCTCATTTGAGTGCTT-----TGGGCCCACTGTA 1622
QY 1518 CAGCAGGCTCTCCCTTGATCAGCCCGCTGTGTAATGACAGAGCCTCCCGCGAGCA 1577
DB 1623 ACAGGGCTATCCCACTCTGAGCCCTGTCTCAGTGAAGTGTGGCGAGCCGGATGGA 1682
QY 1578 CGATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATTAAGCTTCTTACCTAA 1637
DB 1683 CTGCGAG---GAGCCAGACACAGCGCTGCTGCAGAGAGGCTTACTTAAGTCAA 1739
QY 1638 GGAAGTGTCTACCAACCGAGGAACTATTCGAAAGATCTGAAGTTATCTCTCGTGT 1697
DB 1740 AGGATTTCTGCTACAGAAAGAACTATCTCAAGGATTTAAGATTATCCGTTGGTT 1799
QY 1698 TCAGACACAGTGAAGAGAGACGCTATGCGGAAGCACTGAAAGTCTCATATTTCC 1757
DB 1800 CCGCAGCGAGTGTGAAGAGAGCCCATGCTGCGACTGTATGACGTGCTTCTTC 1859
QY 1758 GAATTTTGAACCTTTGCACAAAATTTCACTAAATTTCTCAAGAAATTTGAGCAAGCT 1817
DB 1860 CAACATCGATCCCATCTATGATTTCCAGAGGCTTCTGCGAGGTGAGCAGAGGCT 1919
QY 1818 TGCCCTGTGGAAAGCCGCTCAAAATGCCAAATCAGAG---ATTACCAAGAAATCGCGA 1874
DB 1920 GGCACTCTGGGAAGGCGCTCCCAAGCCCAACAAAGGCACTATCAAGATTCGGGA 1979
QY 1875 TGTCTATGTAAGAACTTCAAGGCGATGAGACCTGCGCGCTCACCTGTGAAGCAG 1934
DB 1980 CATCTGCTCAGGAATGCGGCAATTAAGAGTTTACAGCTTACTTCAAAAGACTGA 2039
QY 1935 CGAGGCTTTGAGAGCCCTGAGATGAATCAAGAGTCCCGCGCTGAGAACTTCTG 1994
DB 2040 CGAGGCTTAAACAGAACTGGAAGAGCTTACCAACGCTGTAAAGATTTGAGGCAAGTGA 2099
QY 1995 CAGAGCTTTGAGCTCAAGAAAGTGTATCTTACCGCTCAACCTTCTCTGCGGCG 2054
DB 2100 CAGAGGTTTGAAGCTGAGAAAGTGTGTACTTGTCTTCAACAGTTCTGTGTAAGCC 2159
QY 2055 ACTGCAACCGGCTCATCTACTACAGCAGGTCTGAGCGGCTGTGCAACACCAACCGCC 2114
DB 2160 CATCAGCGGCTCTGCACTACCGGCTGTGCTGCGCGCTATGCGGACTTATCAGCCC 2219
QY 2115 GAGCCACGCGCACTTCAAGGACTGCGAGCGCTTTGGCAGAGATCAAGAGATGCTGCG 2174
DB 2220 CCGGCACTCATGACTGACTGACCTGACAGCCCTGAAAGCATCAAGAGGTGACAC 2279

QY 2175 ACAGTCCAGGTAGCATGATCAAGATGAGAAATTTCCAGAACTGCAAGAACTCAAGA 2234
DB 2280 CACACTACAGACATTTCTATCTCGGCTGGAGAACTGCAAGAACTTACAGAGTGCAGCG 2339
QY 2235 AGATTTGATTTGCAATTTCTGTTGTTCCGGAGAGGAATTCATCCGTGCGCAG 2294
DB 2340 GAGCTGTGTGGGATGAGAACTTCACTGCTCTGCGAGGAGATTCATCCGTGAGGGCTG 2399
QY 2295 CTTCAAGAGCTTCGGGGAAGGGGCTCCAGCAGCGCAATGTTCTTCTGTTACAGACGT 2354
DB 2400 CTTTCAAGAGCTTCAACAAAGGGCCCTGCAAGAGATGTTTTTCTGTTCAATAT 2459
QY 2355 CTTGCTATACAGAGCGGGGGCTGACGGGCTTCATCACTTAAAGTCCAGCGGAGCT 2414
DB 2460 GTTGTCTACAGAAAGAGTTGACAGAGACACACCTTCCGATCCGGGGCTTCT 2519
QY 2415 CCGCTCTATGTCATGACATTTGAGAGAGCAGAGACAGATGGGGGTGCCCATGCT 2474
DB 2520 TCCCTTCAAGGATCTGTGTGAAAGATGATTAAGATGTTCTTCACTGTTT 2579
QY 2475 GACCTTCGGGGCCAGCGGCACTCATCTGTGCGCGCAATTTCTGCTCGAGATGA 2534
DB 2580 CACCATCTACCGGCTCAGAAACAAATGTTGTGTCAGCGCAGCATCTGGCTGGAAGAA 2639
QY 2535 GAAGTGGTTGAGGACATCAGATGCGCATTTGACTGCGGAGAAAGAGCAGCGCCGC 2594
DB 2640 GAAGTGAATGTCGACTGAATCTCCGATTCAGAGACCAAGATGGCGGTGACACGGC 2639
QY 2595 CCCTGAGTTCTGCGCAGCA-----GCCCTCTGCAACAAAGTCCCTGATGAACAC 2648
DB 2700 CCTGCACTGCAAGCGCGGACTGTGTGCACTGATCCCAAGATCCCAAGAGGTATC 2759
QY 2649 CCGGCTGACAGAGATCAGAGATGACTGAGCGCTTCCGCACTTCGTTGAGAGCGCA 2708
DB 2760 TCTG---GAGCAGAGATCAGAAATGATCTCGGGGTGTCCGAGCTTCCCTGAGGGCA 2816
QY 2709 GGGCCCGCAACCGGCAACAAATGTTGACAGTGTGCTGCAACCGCAACCGAGCTTC 2768
DB 2817 TGCCAGACACCGGCGCAACCAATGACAGTGTGTGTTCTGTACCGAACAACAGCTGT 2876
QY 2769 CATGTGGACTTTCAGATCGCAGTGAAGATCAATGTTGTGAAACCTGAGAAAT 2828
DB 2877 GAGGCAAGCACAAGTGCAGCTGTGAGAAACAGCTTTCAGGATATCTGTAAGAAAT 2936
QY 2829 CAAAAACAGCAACGGGTGCAAGAGTGTGGGTGTGTTCAAAATTTCTGCTGTTCTT 2888
DB 2937 CAAAAACAGTCAATGCTGCAAGAGCTGTGGGTGTGTTTACCAACTTGTGTTGTTCTT 2996
QY 2889 CTACAAATCAACCGAGAAATCATGCTTGCAGGCTGCTGCTGCGGCTACTCGCT 2948
DB 2997 CTACAAACTCATAGAGATGATCAACCACTGGCAGGCTCCGCTGCTGAGCTACAGCT 3056
QY 2949 CACATCCCCCTGTAGTCCGAGAAATCCAGAAAGCTAGTGTTCAGACTGACTTCAA 3008
DB 3057 GAGCATCCCCAGGAGGCGGATGCGATACAAAGACTATGTTTCAACCTCAGTTCAA 3116
QY 3009 GTCCCACTTACTTACTTCAAGCGGAGAAAGCGATACAGTTGAAAGGTGATGGAAT 3068
DB 3117 ATCCCACTTACTTCTTCCGGCTGAGAGCAAGTACAAATTTGAAAGGTGATGAGGT 3176
QY 3069 GATCCGAGTGCACACGACTGTGCTGCGGACGCC 3103
DB 3177 GATCCAGGGGCGCAGACTCAGCGGAGGGGCC 3211

RESULT 12
AB018336
LOCUS AB018336 3997 bp mRNA linear PRI 10-JAN-2004
DEFINITION Homo sapiens mRNA for KIAA0793 protein, partial cds.
ACCESSION AB018336
VERSION AB018336.1 GI:3882306
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1
AUTHORS	Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
TITLE	Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
JOURNAL	DNA Res. 5 (5), 277-286 (1998)
PUBMED	9872452
REFERENCE	2 (bases 1 to 3997)
AUTHORS	Ohara,O., Suyama,M., Nagase,T., Ishikawa,K. and Kikuno,R.
TITLE	Direct Submission
JOURNAL	Submitted (08-OCT-1998) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology, Yana 1532-3, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913, Fax:+81-438-52-3914)
FEATURES	Location/Qualifiers
source	1..3997

/organism="Homo sapiens"
 /mol_type="wRNA"
 /db_xref="taxon:9606"
 /clone="hk05692"
 /sex="male"
 /tissue_type="brain"
 /clone_lib="pbluescriptipti sk plus"
 /dev_stage="adult"
 1..3997
 /gene="KIAA0793"
 <115..3282
 /gene="KIAA0793"
 /note="Start codon is not identified"
 /codon_start=1
 /product="KIAA0793 protein"
 /protein_id="BAA34513.2"
 /db_xref="GI:40788367"
 /translation="RMDEIGCTATVLTQTAGMGAQTPIGVSTTEPMTTILPRMOEKGK
 LHLVGLNDNMELEFDEPKDQVLTITQWKRILNVEDDVGKFGNTQSGVYMKLEK
 MKPIITROIRRRKRVVLELAKFPFPPDPGQIQEYTRYLPALDQKDLLEBTLTADDTT
 AALLTSLQSEIDYDITDEHLKVNELPEQDQKLELLEFHQKVGQTPASDSE
 QVLIARLLEQVIRFPMASDREGTKIQLAVSMGVLYFGTKTLINTFMSKVRKLE
 KRKPKTLHEVYGVQVLPDLEFLKISRECKNFKI CVEYHTFFRLIDQPRPAKAV
 PFSSGSSRSGVGRKQVLDVLPKQSGKIPYRRKIKNTSVRLATDLPKQSIPEF
 EAGTPSPSSPAASVYLSSTLVSGLEPSSSSSLDPOVSVKSAPEAGSGGSS
 AGGDPSPASQGLPRLADPQPKLSTSPQSPSSKSLTSLSPAVQVLPALAGGSS
 LLSSVDSAGAGMDCEPRHKKVPADEAFIYKEILATERYLKOLEYTTWPRSAV
 VKEADMPMTYTLFNSIDYLYFHHGFLAEVQRIALMEGSKATGSHQRIQDILL
 LRRNKQKLPFSTYQRIQDVLVLEKATRCCKLEAVYEFELQVCYLPNTFLKRLK
 IORLLAHLRLTLRLCGYSSGHNDVADCDHALITVEYTLQHLILRELMQKTEIT
 ORDLYVGLNLIAPREYIRIGSCLAHLTKKGLQORMFLPSDMLVYTSKGVAGTSFR
 RGLLPLOGMVYESDNDMSVPHCTYVAAQKTIYVASRLEKEMDNLNSAIDAKK
 SGCTPALPRLERYCTPRPSPNVSLDEQSESDNAGVRSLEHGCGRMLTATHVCK
 LRNTSVSADBSAAVENQSLGLYLLRKPKSHGQKLMVFNFLCYTKHODDYPLAS
 SLPLGASVYIPRADDINDYVFKIQLPSHVVFFPAESKTYFERMWDIVQASSSS
 LAPSIIVDVGDPQSGLEGEMVGRKEE"

ORIGIN

Query Match	33.1%	Score 1037.4;	DB 5;	Length 3997;
Best Local Similarly	60.2%	Pred. No. 1.1e-280;		
Matches 1864; Conservative	0;	Mismatches 1186;	Indels 45;	Gaps 7;

QY 21 GCCGACCCAGAGATCA CGACTGGGGGGCCCCGGAAAAATTGGGGAGTCA GTACTTTGGAAG 80

Db 150 GCAACACTGCAGGGATGCGCTTGGGGGCCCAACCCCTCTGGGAGATTACACCCCTTGAGCC 209

QY 81 TGGACAGAGGCGGCCCAACACCTTACGAAAACTGTCATCAATCAATCCAGATGCT 140

Db 210 TGGGGAAGCTCTTTGGCCAGAAAGCAAGAGACACTGCACCTCAAGATGAAGCTGCT 269

QY 141 GGATGACACCCAGAGAGCATTTGAAGTTCACAAAGAGCTCCTGGGAAGGTGCTCTGGA 200

Db	270	GGAGACACACATGGAATATTTGACATTGAGCCTAATATGGATGCGCAGGATTAATCAGAC	329
Qy	201	TGCAGTTTGCACCACTCAACTCGTGGAAGGTGACTATTTTGGCCTCGAGTTTCTGTA	260
Db	330	ACAAGTGTGGAAAGGTTTAAACCTGGTGAATGATGACTTCCGGATGGAAGTTTCAAAA	389
Qy	261	TCACAAAAAGATCACGGTGTGCTGATCTCTTAAACCCATTGTGAAACAGATTAAAG	320
Db	390	TACTCAGTCTACTGGAATTTGGCTTGAACCTATGAAACCATCTATTAGCGAAATACGAA	449
Qy	321	GCCAAAGCAGTGTGTTGTTAAGTTTGTGGTAAATTTCTTCCGCTTGAACAACAAC	380
Db	450	GCCAAAGAAATGTGCTGCTCGCCTTAAGCTGTAAATTTTTCACCTCATCTGTGTGAGCT	509
Qy	381	CCAAAGAAATCAACAAGGTACCGTTCCGGCTGACGTGAAGCAGGACTTTGGCTCAAG	440
Db	510	ACAAAGAAATATACAAAGATCTTGTTGCTTGCACTTAAGAGAGACCTGCTGGAAHA	569
Qy	441	CAGGTTGACGTGTAAATGACAACAAGCAGCTCTCTGATTTTCAACATTTGTGCATCTGA	500
Db	570	GCGTTTGAACGTGTGCTGACACACAGGCGCCCTTCTCACGTCCTCACTCTGACGTGGA	629
Qy	501	GATTGGGGAATTTTATGAGCCTTTGGAACAAGACACTTAAAGCAAAAAATAATATCATCC	560
Db	630	AATGAGGATTAACATGAAACCGCTGGAACGAGAGCACTCAAAAGTGAACGATATTTGCC	689
Qy	561	TCACAAAGACGCACTAAGGACAAATTCGTGAATTTCAACATTAACACATTTGGACAAAC	620
Db	690	TGGCCACAGACACTGCTTGTAAGAAATCTGAATTTCCATCAAGAGACGCTGGGCCACAC	749
Qy	621	ACCACGAGATCAGATTTCCAGCTCTTAAGATTTGCCGTGGCTAGAGATGATGGAAT	680
Db	750	ACCTGCTGAGTCGATTTTCCAGGTGCTCGAAATTTGCTCGAAAGTTGGAATGTACGGCAT	809
Qy	681	CCGGTTGACCCCGGCCAAGGACAGGAAAGCAGAAATCAATCTGGCCGTGGCAACAC	740
Db	810	CAGATTTCACTGGCTTCTGACAGGAAAGAACCAAGATTTCAACTGGCAGTTTCCCAT	869
Qy	741	GGGAATTCAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCCAAGTGC	800
Db	870	GGGTGTACTCGTGTTCAGGGGACCAACAAATCAACATTTCACTGTGTCMAAGTCCG	929
Qy	801	GAACTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGCGCANATGCCATATATGC	860
Db	930	TAAACTAAGCTTCAAGAGAAAGAAATTTCTATCAAACTTATCAAGAGTTCATGAGCC	989
Qy	861	GTACCAAGATCTCTGGAATTTCCGATGAGCGCAGTCGGAATTTGCAAGTCTTCTGAA	920
Db	990	TTACCAAGGACATTAAGATTTTGTGGGATGAGAGATGAATGTAAAGAACTTCTGAA	1049
Qy	921	AATCTGTGTGAACATCATGCTTCTTTAAGCTTTTGAAGAGCCCAACCAAGCCCA	980
Db	1050	GATTTGTGTGAGATCAACCTTTTAAACCTTTTGAACCACTTAAGCCAAAGCAAA	1109
Qy	981	GCCGCTCCTCTTAAAGCCGGGGGTCACTATTCGGTTCAAGTGTGAGTCTGAAAGCAGGT	1040
Db	1110	AGCGCTCTTCTTCAAGCCGGGCTCTCTCTTCAATACATGAGAGAACTCAAGAAACA	1169
Qy	1041	TCTGACTATGTAAAGAGAGACATTAAGAAAGTCAAGTTTGAAGAGACACAGCA	1100
Db	1170	AGTAGATTAATTTCAAGACAGTGAATGAAGAAATTCATATGAAGAAAGGCACACAA	1229
Qy	1101	GATTCAATTCATCCGAGCCTTGTTCAAGCCTCAAGAACTGAATTCGGAAGTGTGGA	1160
Db	1230	GACCCACAGTCC-----GTTGAGCTCTGACTGACAGCTCAACAA	1271
Qy	1161	GCACTTCAGACAGACCAAGCCTTAATTTGGAGAAAGTGGCCGAATCTCAGGGGGCCA	1220
Db	1272	ACAAAGATCTCAATCCCGAGGAGATGAAGACCTCGCTCCCACTTCAAGGAATGC	1331
Qy	1221	GAGCTCCCGGAGAAAGAAACCGAAGGTTTCCGCGGGGAGCTGGGGTTCGACCCGAG	1280
Db	1332	CTTTTACTGTGCTCTCTCCCTCACTCTGTGTCCTCTCGGCTGCAAGATTTAAAGACAG	1391

QY 1281 CCCTGGCGGAGGAGAGCCCGGGGTAAACAAGAGCGGAGCGGCCCTCGGGCC 1340
 Db 1392 CAGACAGCTCTCTACAGATCCCAAGGTTCTTACGTCAAGAGTCCAGCTGCAGAGAGCG 1451
 QY 1341 CACGGAAGAAGAGAGAGG---TGGTTAAGATAGAACCCAGAGATTAACCTGAGCC 1397
 Db 1452 CAGTGAAGAGTGGCTGAGAGGCCCCGACACATGGGCCACCCCTCGGGCCCCCGC 1511
 QY 1398 CCGGCAAGCAAGAGCTCCCTGACGTGAGTCTTACCTTTCGAGCTGTGTGA 1457
 Db 1512 ACTCCAGCTGTGTGAGGCTTTTCAAGAGATCCAGCTTCTCCCTCAGCGGAA 1571
 QY 1458 CTGCGAGGGGGAGTGGCCCTGCGCAAGTACCTTGTCTCCACCTGAGCCCGACAC 1517
 Db 1572 GAGCCCCCTGAGTCTAGAGCCCTGATTTTCAAGTGCTT-----TGGGCTCAGCTGA 1622
 QY 1518 CAGCAGAGCTCTCCCTTGTATCAGCCGCTGTGAATGACAGAGCTGCCCCGAGCGA 1577
 Db 1623 AAGAGGCTATCCCACTCCTGAGCCCTGTCTCAGAGATGCTGGGAGCCGGAGTGA 1682
 QY 1578 CGATGAGATGAGGGCCGAGAGAGATTTCCCACTGATTAAGCTTACTTCAATGCTAA 1637
 Db 1683 CTGCGAG---GAGCCGAGACAAAGCGCTGCTGAGAGAGAGGCTTACTTCAATGCTAA 1739
 QY 1638 GGAAGGTCTACACAGAGGAAATATGAGAGATCTCGAAGTTATCATCTTCGTGTT 1697
 Db 1740 AAGGATTTCTGCTACAGAGAAATATCTCAAGAGATTTAAGATTTATTCAGTGTGTT 1799
 QY 1698 TCAGAGCAGATGAG 1757
 Db 1800 CGGAGCGAGTGTGTAAG 1859
 QY 1758 GAATTTTGAACCTTTGACAAATTTTCACTAATTTTTCAGAGAAATTTGAGAGAGAG 1817
 Db 1860 CAAATCGATCCCATATGAGTTTCCAGAGGCTTCTGCGAGAGTGAAGAGAGAGCT 1919
 QY 1818 TGGCCGTGAG 1874
 Db 1920 GGCACTCTGAG 1979
 QY 1875 TGTGATGCTGAAGAGATTTCAAGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1934
 Db 1980 CATCTGCTCAGAGAGATGCGCAAGTTAAAGAGTTTACAGTACTTCCAAAGAGATGA 2039
 QY 1935 CGAGGCTTTGAG 1994
 Db 2040 CGAGGCTTTAAG 2099
 QY 1995 CAGAGCTTTGAG 2054
 Db 2100 CAGAGGTTTGAAG 2159
 QY 2055 ACTGCAACCGGCTATGACATCAAGAGAGGCTCTGAGAGGCTGTGAAAACCAACCGCC 2114
 Db 2160 CATCCAGCGGCTGTCTGACATCAAGGCTGTCTGCTGCGCCGCTATGAGGACATTAAGCC 2219
 QY 2115 GAGCAGCGGCTTCAAG 2174
 Db 2220 CGGGCAGCATGACTGAGTGCATGACATGACGCGCTGAAGGATCAGAGAGGTGACAC 2279
 QY 2175 ACAGCTCCACGAGTACATGATCAAGATGAGAAATTTCCAGAGAGTCAAGAACTCAAGAA 2234
 Db 2280 CACACATCAAGACATTTCTATCTCGGCTGAGAACTGCAAGAGCTAAGAGTGAAGCG 2339
 QY 2235 AGATTGATTTGATGACATCTTGTGTTCCGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 2294
 Db 2340 GGAAGCTGTGGGATGAGAACTTATGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2399
 QY 2295 CCTCAG 2354
 Db 2400 CCTTCAAGAGTCAACAAAG 2459

QY 2355 CTGTATATACAGAGCGGGGGCTGACGGCTCTCAATCAGTTTAAGTCCAGCGGAGCT 2414
 Db 2460 GTTGTCTGTACAAAGAAAGAGTTTCAAGAGACACACTTCCGATTCGGGGCTTCT 2519
 QY 2415 CCGGCTCTATGAGATGACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2474
 Db 2520 TCCCTTCAAGAGATCTGTGTGAAGAAAGTGAATACAGATGTGTCTTCAACATGTTT 2579
 QY 2475 GACCTTCCGGGGCAGCGGAGCTTCATCTGTGCGGCAAGTTCTCGTCCAGATGA 2534
 Db 2580 CACCATCTACGGGCTCAAAAAAATGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAA 2639
 QY 2535 GAAGTGGTTGAGAGATCAATGAGTGAATGCACTTGAGCGAGAGAGAGAGAGAGAGAG 2594
 Db 2640 GAGTGAATGCTGAGACTTAATCTCCGATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 2699
 QY 2595 CCCTGAGTCTGTGAG 2648
 Db 2700 CCTGCACTGCCAGGCGGAGCTGTGTGCACTGTGCTCCCAAGATCCCAAGAGAGATC 2759
 QY 2649 CGGAGCTGACAGAGATCAGAGATGACCTGAGCGCTCGCGACATCGCTGAGAGCGCA 2708
 Db 2760 TCTG---GAGCAGAGATCAGAGATGATCTGGGGGTGTCCGAGCTCCCTGAGAGGCA 2816
 QY 2709 GGGCCCGCAGCGGAG 2768
 Db 2817 TGGCCAGACACGGGAG 2876
 QY 2769 CATGTGAGCTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2828
 Db 2877 CAGGAGAGACAG 2936
 QY 2829 CAAAAACAGAAAGGCTGAG 2888
 Db 2937 CAAAAACAGTCAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2996
 QY 2889 CTACAAATCAGACAG 2948
 Db 2997 CTACAAATCAGACAG 3056
 QY 2949 CACATCCCCCTGTGAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3008
 Db 3057 GAGCATCCCAAG 3116
 QY 3009 GTCCCACTCTACTACTTCAAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3068
 Db 3117 ATCCCACTCTACTTCTTCCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3176
 QY 3069 GATCCGAGTGCACACAGCTGTGCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3103
 Db 3177 GATCCAGGGAG 3211

RESULT 13
 BC009153 3719 bp mRNA linear ROD 29-JUN-2004
 LOCUS Mus musculus PERM, RhoGAP and pleckstrin domain protein 2, mRNA
 DEFINITION (cdna clone MGC:304 IMAGE:2655209), complete cds.
 ACCESSION BC009153
 VERSION BC009153.1 GI:14318718
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 3719)
 Strusberg R.L., Peltingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Scheffler C.F., Bat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,

Scheetz, T. E., Brownstein, M. J., Usdin, T. B., Toshiyuki, S., Carinini, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J., Abramson, R. D., Mullahy, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W., Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakeley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmitz, J., Myers, R. M., Butterfield, Y. S., Krzywinski, M. I., Skalska, U., Smalins, D. E., Scherneck, A., Schein, J. E., Jones, S. J., and Marra, M. A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

JOURNAL

2 (baeas 1 to 3719)

Strausberg, R.

Direct Submission

Submitted (05-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LNL)

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P. H., Garcia, A. M., Lu, X., Hulyk, S. W., Louesged, H., Kowis, C. R., Sneed, A. J., Martin, R. G., Muzny, D. M., Navarat, A. N., Gibbs, R. A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: IRAX Plate: 5 Row: m Column: 20

This clone was rejected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21704021.

Location/Qualifiers

1. 3719

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N-3"

/db_xref="taxon:10090"

/clone="MGC:6304 IMAGE:2655209"

/tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."

/clone_11b="NCI CGAP Mam2"

/lab_host="DH10B"

/note="Vector: pcwv-sport6"

1. 3719

/gene="Farp2"

/note="synonym: Fir"

/db_xref="GeneID:227377"

/db_xref="MGI:2385126"

99. 3296

/gene="Farp2"

/codon_start=1

/product="ERM, Rhogef and plectestrin domain protein 2"

/protein_id="AAH09153.1"

/db_xref="GI:14318719"

/db_xref="GeneID:227377"

/db_xref="MGI:2385126"

/translaction="MGEIRGTALPTSGTGLGOTALIGVSTLEBEOISLSPMOEKHM RTRVLDLSTVELFDIEBKCDGVLITVYKHLNIBCDYFGLBKPKNVOSYIMLEPM KRIIOVRKRNAAVIRLAKFPBPBGOLQEBYTYLTFALQKDLLEBRLICTANTA ALLISHLQSEIGDDEITDRHLKNAEYLPHOESKLEKIDFHQHHTGQTAEEDFO VLEIARKLEMYGIRFPMASDRGCTKINILAVSHGVLVFQGTIKINTFMNSKVRKLSFK

ORIGIN

Query Match 32.3% Score 1013.2; DB: 6; Length 3719;

Best Local Similarity 59.3%; Pred. No. 7.4e-274;

Matches 1846; Conservative 0; Mismatches 1238; Indels 30; Gaps 6;

22 CCGACCCGAGATCAAGACTGGGGGCCCGGAAATCGGGATCAGTACCTTGGAACT 81

132 CCAACTTCAAGGACACGCTGGGTGGCCAGACCCGATTTGAGTACCTTGGAGCA 191

82 GGACGAAGCCGCCCAACACTTCAAGAAATCTGTCTCATCAAAATCCAGATGCT 141

192 GAGCAGACCTGTCAACCCAGATGAGAGAACATGCGCATCAGATGAAGCTGCTG 251

142 GATGACACCCAGAGGACATTTGAAGTTCCAGAAAGCTCCGTGGAGAGTGTCTGAT 201

252 GACAGACTGTAGATTTATTTGACATCAAGCCGAAGTGTGACGGGCAAGTCTTACTG 311

202 GCAATTTCAACCACTCACTCACTGAGAGGTGACTATTTTGGCTCGAGTTCTGAT 261

312 CAAGTGTGAAGCATTTGATGATCTCATGATGTGACTATTTGTGTGAGTTCAAGAT 371

262 CACAAAAGATCAAGGTGTGCTGATCTCTAAACCATTTGTAACATTTGAAGG 321

372 GTCCAGTCTCACTGATTTGGCTTGAACCTATGAACCATATTTAGCAAGTGAAG 431

322 CCAAGACAGTTGTGTTAAGTTGTGTAATCTTTCGCGCTGACACACACACACTC 381

432 CCAAAAATGGGGTCTTCCGCTGACATTAATTTTCCGCTGACCTGCTGACGCTG 491

382 CAGAAGAACTCAAGAGTACCTGCTCGCTGACAGTGAAGCAGAGACTTGGCTCAAG 441

492 CAGAAGAGTACCAAGGATCTGTTTGCCTTGAACCTCAAGAGAGACCTCTGAGGAA 551

442 AGGTTGACGTGTATGACACAGCGAGCTCTTGAATTTACACATTTGTCATGTAG 501

552 CGCTGACCTGACCGGCAACCTGACAGCCCTTCTCATATCCACCTTCTGCAAGTC 611

502 ATTGGGATTTGTAAGAACCTTGGACAGAGCACTTAAGCAAAATTAATACATCT 561

612 ATCGAGATTTGATGAAGAACCTGATGAGAACCTCAAGAACCAATGAATACCTG 671

562 CAGCAAGACGACTAGAGGACAAATCGTGAATTTCAACCAATTCAGACAAACA 621

672 AACGAGGAAATCTCTGAAGAAATACATGACTTCAACAGAGGACACGGGCCAGAT 731

622 CAGCAGAAATCAGATTTCCAGCTCTTGAAGATTTGCGCTGAGTGAAGTGTAGATC 681

732 CCTGACAGATGAGATTTCCAGCTGCTTGAATTCAGAAAGCTTGAATGTATGCAATC 791

682 CGCTTGACCCGCGCAAGAGAGGACGGAAGATCAATCTGCGCTGCAAGACAG 741

792 AGTTTCAACATGCTTCTGACAGAGAGGACCAAGATTAATCTGACAGTTTCTCAATG 851

742 GGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCACTGCGGCAAGGTGCG 801

852 GGTGTCTGTGTTCAGAGGTACACCAAAATCAATACCTTCAATGCTGCAAAAGTCGT 911

802 AAGTACCTTCAAGAGAGAGCGCTTCTCATCAAGCTTCGCGGACAGATGCAATGTCG 861

Db 912 AAACTAAGCTTCAAGAGAAAAGTTTCTTATCAAACTCCACCCTGAGGCTCATGAGCC 971
Qy 862 TACCAAGATTAATCTTGAATTCCTGATGCGAGTCGGATTTTCGCAAGTCTTCTGAAA 921
Db 972 TACCAAGACAAGTTAGATTTCTTGGGATGAGAGATGAATGTAAAGAACTTCTGAA 1031
Qy 922 ATCTGTGTGAACATCATGCTCTTCTTGAAGTTTGAAGAGCCCAACCAAGCCCAAG 981
Db 1032 ATATGTGTGAGACATCACTCTTTTAACTCTGACAGGCTTAAGCCAAAGCAAG 1091
Qy 982 CCGGCTCTTTAAGCCGGGGGTCATATTCGGTTCAAGTGTGGAATCAAGAGAGTT 1041
Db 1092 GCTGTCTTCTTACGCGAGGCTCTCTCTTCAATACAGTGAAGAACTCAGAAACAATA 1151
Qy 1042 CTCGACTATGTTAAAGAGAGACATTAAGAGTGCAGTTTAAAGAAAGCAGACAG 1101
Db 1152 GTAGATTATGTCAAAACGGTGAATGAAGAAATTCATACAGAAAGACGGCAGATTAG 1211
Qy 1102 ATTC--ATTCTATCCGAGCCTTGTCTTCAAGCTTACAGAACTGAATTCCGAAGTCTG 1158
Db 1212 ACTCGACATCTTCTTATGCTGTGACTGTATGATCTGCTTAAACAGAGCTCTCTTACC 1271
Qy 1159 GAGCAGTCTCAGCAGACACAGCCTTACATTTGAGAAAGTCCGAAATCTCCAGGGG 1218
Db 1272 GATGGCTTGAAGACTTCTGCTCTCTGTCTTCAAGAAATGTCTCTTTTATTC----- 1324
Qy 1219 CAGAGTCCCGGAGAGAAAGAAACCGAAGTTTCCGCGGGAGACCGGGGTGACCCG 1278
Db 1325 ---ACCCCTAGTCTTCTTGTCTCTCTCTGCTGCTCCCAATTTGAAGACAGCAGAG 1381
Qy 1279 AGCCCTGCGCCGAGAGAGAGCCCGCGGGTAAACAAGAGCGGACGGAGCGCTCGAGG 1338
Db 1382 CTCCTTGTGATCCGAGGCTCCGTCATCAAGAGACAGCAGCAGAGAGAGAGCGG 1441
Qy 1339 CCCACGAGAGAGAGAGAGATGCTTAAAGATAGACCCAGCAGAGTAACTCAGACC 1398
Db 1442 ACCA--TCAATCAATGACCGGCCACAGACACAGTCGGCCATCTCCCTGAGCCCCCG 1499
Qy 1399 CCGCAGCCAAAGCAGAGCTCCCTGACTGGCAGTCTTCACTTTCCAGGCTGTGTGAAC 1458
Db 1500 CTCGCGCTGTGTCAAGGCTTTTCTATGATGATGCTTCAAGCTTCTCCCTCAAGCC 1559
Qy 1459 TGGCAGGGGGAGTGGCCCTGCAACGTGACCTTGTCTCCAACTGAGCCCGACACC 1518
Db 1560 AGCCACTGAGCTGTGCTGCTGAGCTTCAAGCCGCG-----CTGAGTACAGCTGAG 1610
Qy 1519 AAGCAGGCTCTTCCCTTGAATCAGCCGCTGTAATGACAGGCTGCCCCGAGCAGAC 1578
Db 1611 CAGGGTGATCCCTGTGTCTGAGCCCTGTCTCAAGTGGCCGTGTACAGCAGAGATGAC 1670
Qy 1579 GATGAGATGAGGGCCGAGAGAAAGATTCCTCACTGATTAAGCGTACTTCAATAGTAA 1638
Db 1671 AACCAAGAGAGAGAGAAACAAGCACAATGCCAGAAATGAGGCTATTTCAATAGCAAG 1730
Qy 1639 GAAGGTCTAACACCGAGGGAATATCTGAAGAGATCTGGAAGTATCTTCCTGTGTGTT 1698
Db 1731 GGAATTTCTGCTACAAACGAACTTATCTGAAGATTTAGAAATATCACTGTGTGTTC 1790
Qy 1699 CAGAGCAGATGAGCAAGAGAGAGCAGCATGCGGAAAGCACTGAAAGTCTCATATTTCCG 1758
Db 1791 CGAGGTGTGTATCAAGAGAGAGGCAATGCTGAGCCCTGATGAGCCCTGCTTTCTCC 1850
Qy 1759 AATTTGAACCTTTGACAAATTTCACTAATTTTCTCAAGAAATTTGAGACACACTT 1818
Db 1851 AACATTTATCTCAAGTCTTACAGATTCACAGAGCTTCTTCAAGAGTGAACAGAGCTG 1910
Qy 1819 GCCCTGTGAGAGAGCGCTCAATGCGCA---AATCAGATTAACCAAGAAATCGCGCAT 1875
Db 1911 GCACTTGGAGAGGCGCTTCAAGTGCACCTTAAAGGTATATCAAGCAGAAATCGGGAGAC 1970
Qy 1876 GTCATGCTGAAGAACTTCAAGGAGTGAAGCAGCTGGCGGCTCACTGTGGAAGCAGC 1935

Db 1971 ATCTCTCAGAGAACATGCGTCAGTTAAAGAAATTTACTAGCTACTTCCAAAGACAGAT 2030
Qy 1936 GAGGCTTGAAGGCGCTTGAGAGATGGAATCAAGAGCTCCCGCGGCTGAGAGACTTCTGC 1995
Db 2031 GAGGTCTTAACAACTGGAAAGAGCCACAAACACTGTAAAGATTTGAGGAGTCTTAC 2090
Qy 1996 AGAGACTTTGAGCTGAGAAAGGTGTGTTACTACCGCTCAACACTTCTCTGCGGCA 2055
Db 2091 AAAGATTTGAGTCCAAAGAGTCTGTCTCTCTCTCAACACTTCTCTGAGAGCC 2150
Qy 2056 CTGCACCGCTATGACATTAACAGAGTCTTGAAGCGGTGTGCAAAACACACCGCGG 2115
Db 2151 GTCCAAAGGCTAGTCCACTACCGTGTGTGAGCGGCTGTGTCTCACTCTCTCT 2210
Qy 2116 AGCCAGCGGACTTCAAGGACCTGCGCAGCGCTTGTGAGAGATCAAGGATGTGGCA 2175
Db 2211 GGGCAGCGGACTATGCGCACTGCATGAGGCACTAAGGCCATCAAGAGTACACACC 2270
Qy 2176 CAGCTCAGGTTACATGATCAAGATGAGAAATTTCCAGAAAGCTGCACGAACTCAAGAA 2235
Db 2271 GAGCTCAGCAAAAGCTTACCGGCTGGAAGAACTTCAAGAAATTTGACGAGTCAAGCA 2330
Qy 2236 GATTGATTTGCAATGACAAATCTTGTGTTCCGGAGAGGAGTTCAATCCGTCTGGCAGC 2295
Db 2331 GACTGTGTGTTGTAAGAAACCTCATTTGCTCTGGAGGAGATTATCCGTGAGGGCTGC 2390
Qy 2296 CTCAGAAAGCTCTCGGGAGAGAGGGCTCAGAGCGGATGTTCTTCTCTTCAACGACGT 2355
Db 2391 CTGCACAACTCAACAAAGAGGCTGCAACAGAGAGATTTTCTGTCTCAGATAG 2450
Qy 2356 CTGCTATACAGAGCGGGGGCTGACGGCTCCAACTGATTTAAAGTCCACGGGAGCTC 2415
Db 2451 TTACTGTATACAAAGAAAGTGTCAAGAGACCAAGTCTTCCGATCCGTGGCTTCTT 2510
Qy 2416 CCGCTCTATGCAATGCAATTAAGAGAGCGAAGCAGATGAGGGGTGCCCATCTGCTG 2475
Db 2511 CCACTCGTGGCATGTGTGTAAGAAAGTGAAGATGAGTGTCTTCTTCAATGCTTC 2570
Qy 2476 ACCCTCGGGGCGAGGCGAGTCCATCATGTGGCCGCAATTCGTGCTCCAGATGAG 2535
Db 2571 ACCATTTATGAGCTCAGAAACAAATGTGTAGAGCCAGCATGCGCTAGAAAGAA 2630
Qy 2536 AAGTGGTTGAGGACATCCAGATGCGATTTGACCTTGGCGAGAGAGAGAGCCCGGCT 2595
Db 2631 AAGTGAATGAGGACCTGTAATGACAGATTCAGAGCCAAAGCTATTCGATCTACCC 2690
Qy 2691 CCGTGAATCTGCGCAGCAGCCCGCTGCAACAAAGTCCCTGATGAGCCACCGGGCT 2655
Db 2691 CAGGTGCTGTG--GAGGCGGGGTATACGTGTAACCTTGATCTTCTGATGAATTC 2747
Qy 2656 GACAGAGATCAGAGATGACCTGAGCGCTGCGCAATCGCTGAGAGCGCCAGCCCG 2715
Db 2748 TCTCTGGAAGAAATCAGAAATGTGTGAGAGAAACCGGGGCTCCCTGAGGGGAAACAGCCAG 2807
Qy 2716 CACCGCGCAACAATGAGTGCAGTGTGTGCAACCGCAACCGAGCTTCTCATGTG 2775
Db 2808 CACCGGCAATTAACAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2867
Qy 2776 GACTTCAGATGCAAGTGAAGATCAATGTGTGAAACCTGTGTGAAATTTCAAAAC 2835
Db 2868 GACACAGTGAAGCTGTGTGAACCAAGCTTTCAGATATCTGTGTGAAGATTTCAAGAAC 2927
Qy 2836 AGCAAGGTTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2895
Db 2928 AGTAATGTGTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2987
Qy 2896 TCACACAGAGCAATCATCCCTTGCAGCCTGCTCTGCTGCTGCTGCTGCTGCTGCT 2955
Db 2988 ACACATCAGAGATGATCAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3047
Qy 2956 CCCTGTAGTCCGAGAACTCAGAAAGACTACGTGTCAAGCTGACTTCAAGTCCAC 3015
Db 3048 CCAGGAGGCTGACAGATCAACAAAGACTATGTCTTCAAGCTCAATTTCAATCTCAC 3107

	3016	GTCACACTCATTCAAGGGCGGAAAGAGATACACGTTCCGAAGAAGTGAAGTAAGTATCGC	3075
Db	3108	GTCACCTCTTTCGCGGTGAGACGACATGCACATTGTAAGTGATGAGCATCATCAA	3167
QY	3076	AGTGCACCAAGCTCTGCTCGGACGCCACGTTGAGCCACAAGAAGTCTCTT	3129
Db	3168	AGGGCCAGCAGCTCACCGAGGAGACCCTCAAGTTTCACTCAGACCTGCTCACAT	3221
RESULT_14			
LOCUS	BC027077		
DEFINITION	BC027077	2431 bp	mRNA linear ROD 21-Oct-2003
	Mus musculus FERM, RhogEF (Artygef) and plectestin domain protein 1		
	(chondrocyte-derived), mRNA (cDNA clone IMAGE:5010682), partial		
	cds.		
ACCESSION	BC027077		
VERSION	BC027077.1	GI:20071584	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;		
	Sciurognathia; Muridae; Muridae; Mus.		
REFERENCE	Krauseberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,		
AUTHORS	Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,		
	Altschul,S.F., Zeeberg,B., Buelow,K.H., Schaefer,C.F., Bhat,N.K.,		
	Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,		
	Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.W., Hong,L.,		
	Scapellato,M., Soares,M.B., Bonaldi,M.F., Casavant,T.U.,		
	Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toobylyk,I.S.,		
	Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.T.,		
	Abrahamson,R.D., Muliahy,S.J., Bosak,S.A., McGowan,P.J.,		
	McKernan,K.U., Malek,J.A., Gunaratne,P.H., Richards,S.,		
	Worley,K.C., Hale,S., Garcia,A.W., Gay,L.U., Hulyk,S.W.,		
	Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,		
	Fahney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,		
	Sanchez,R., Whiting,M., Madan,A., Young,A.C., Shcherbak,Y.,		
	Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,		
	Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,		
	Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailis,D.E.,		
	Schneker,A., Schein,J.E., Jones,S.J. and Maric,M.A.		
	Generation and initial analysis of more than 15,000 full-length		
	human and mouse cDNA sequences		
TITLE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
JOURNAL	12477932		
PUBMED	2 (bases 1 to 2431)		
REFERENCE	Strauseberg,R.		
AUTHORS	Direct Submission		
TITLE	Submitted (04-APR-2002) National Institutes of Health, Mammalian		
JOURNAL	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk		
	Email: cgabs-xemail.nih.gov		
	Tissue Procurement: Gilbert Smith, Ph.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA library Arrayed by: The I.M.A.G.E. Consortium (ILN)		
	DNA Sequencing by: Baylor College of Medicine Human Genome		
	Sequencing Center		
	Center code: BCM-HGSC		
	Web site: http://www.hgsc.bcm.tmc.edu/cdna/		
	Contact: amg@bcm.tmc.edu		
	Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisgeed, H.,		
	Kowitz, C.R., Sneed, A.J., Martin, R.G., Muzny, D.W., Nanavati,		
	A.N., Gibbs, R.A.		
	Clone distribution: MGC clone distribution information can be found		
	through the I.M.A.G.E. Consortium/ILN at: http://image.llnl.gov		
	Series: IRAK Plate: 45 Row: c Column: 12		
	This clone was selected for full length sequencing because it		

```

passed the following selection criteria: Hexamer frequency ORF
analysis.
FEATURES
Location/Qualifiers
1. .2431
source

```

gene

CDS

misc_feature

misc_feature

ORIGIN

Query	2074	TACAGAGAGTCTCGAGACGCGCTGTGCAAAACCAACCGCGGAGACGACGCGGACTTACAG	2133
1	TATAGAGATGTCCTGAGAGAGGCTGTGCAAGACCAACCAACCAAGCGGACTTACAG	60	
Query Match	28.0%;	Score 878.6;	DB 6; Length 2431;
Best Local Similarity	89.3%;	Pred. No. 7e-36;	
Matches 947; Conservative	0;	Mismatches 114;	Indels 0; Gaps 0;

OY		2134	GACCTGCCAGCCGCTTTTGGCAGAAGATCAACGGAGATGTGTGCACAAGCTCCAAGGTACATG	2193
Db		61	GACTGCAGAGCTGCCTCGCCGGAGATCAACGAAGATGTGTGCACAAGCTCATGTATCAATG	120
OY		2194	ATCAAGATGAGAAATTTCACAGAGCTGCACGAACCTCAAGAAAAGATTGATGGCATTTAC	2253
Db		121	ATCAAGATGAGAAATTTCAGAAAGCTGCATAGCTCAAGAAAAGATTGATGGCATTTAC	180
OY		2254	AATCTTGCTGCTCCGGGAAGGAGTTCAATCCGTCTGGGCAGCTTCAGCAAGCTCTCGGG	2313
Db		181	AATCTTGCTGACCCCCAGGAAGGAGTTCAATCCGCTCTGGGCAGCTTCAGCAAGCTCTCGGG	240
OY		2314	AAGGGGCTTCACAGACGCGATGTTCTTCGTTCACAGAGCTCCGTCTTAACAAGCCGG	2373
Db		241	AAGGGGCTTCACAGACGCGATGTTCTTCGTTCACAGAGCTCCGTCTTAACAAGCCGG	300
OY		2374	GGGCTGACGGCCTCCCAATCAGTTTAAAGTCCACGGGCAGCTCCGCTCTATGACATGACG	2433
Db		301	GGACTGACGGCATCTTAATCAGTTTTAAGTCCACGGGCAGCTCCCACTTATGACATGACG	360
OY		2434	ATTGAGGAGAGCGAAGACGAGTGGGGGGTCCCCTGCTGACCCTTCGGGGCCAGCG	2493
Db		361	ATTGAGGAGAGTGAAGAGAGAGTGGGGTGGCCCCCACTGCTGACCCTTCGGGGCCAGCG	420
OY		2494	CAGTCCATCATCTGAGCGCCGACGTTCTGAGTCCGATGAGAAAGTGGGTTGAGGACATC	2553
Db		421	CAATCTATCATCTGAGCTGCGACGTTCCGGCTCAGAGATGAAAAATGATGAGAGCATTC	480
OY		2554	CAGATGGCATTGACTGTGGCGAGAAAGACAGACGCCGCCGCTGAGTTCTTGCCGACG	2613
Db		481	CAGATGGCCATTGACTGTGGCAGAAAGAACATGAGCCACCCCGAGCTCTTGCCGACG	540
OY		2614	AGCCCCCTTGACAACATGCCCTGTGATGAACCAACCGGGCTGACAGAGATCAGAGAT	2673
Db		541	AGCCCCCTTGACAACAGTCCCCGGATMAACCAAGGGGTGACAGAGATCAGAGAT	600
OY		2674	GACCTGAGCGGCTCGCGCACATCGCTGAGAGCGCCAGGCCGCCGACCGGGCAACAATG	2733
Db		601	GACCTGAGTGGCTCTCGGACATCGCTGAGAGCGCCAGGCCCTTCACCGTGGCAACATG	660
OY		2734	GTGCACGTGTCTGTGCACCGCAACACAGCGTCTTCATGTGTGAGACTTCAGCATCGCATG	2793
Db		661	GTGCACGTGTCTGTGCATCGCAGACACAGTGTCTTCATGTGTGAGACTTCAGCATGTGTG	720
OY		2794	GAGAAATAGTTGTCTGGAACCTGTGTGAAATTCABAAACACACAAACGGTGGCAGAG	2853
Db		721	GAGAAACAGCTCTCTGGGAATTGTGTGAAAGTTCAAGAACACACAAATGTGTGCAGAG	780
OY		2854	CTGTGGGTGTGTTTCACAAATTCTGTGCTGTCTTCTTACAAATGACACAGGACAAATCAT	2913
Db		781	CTGTGGGTGTGTTTCACCAACTTCTGTGCTGTCTTCTTACAAATGACACAGGACAGTCAAC	840
OY		2914	CCCTTGGCAGCGCTGCTGTGTGTGGGCTACTGTGCTCACCATCCCTCTGAGTCCGAGAAC	2973
Db		841	CCCTTGGCAGCGCTGCTGTGTGTGGGCTACTGTGCTCACCATCCCTCTGAGTCTGAGAAC	900
OY		2974	ATCCAGAAAGCTACGTGTCAAGCTGACATTCAAAGTCCACAGCTCTACTATTAGGGCG	3033
Db		901	ATCCAGAAAGCTATGTGTGTCAAGCTGACATTCAAAGTCCACAGCTCTACTATTAGGGCT	960
OY		3034	GAAAGCGAGTACAGGTTGAAAGGTGTGATGATGATCCGACAGTGCACCCAGCTTGCC	3093
Db		961	GAAAGCGAATACATTTGAAAGGTGTGATGATGATCCGAAAGTGCACACAGCTTAGCC	1020
OY		3094	TTCGGAGCCCACTGTTGAGCCCAACAAAGTCTCTTGTTGA 3134	
Db		1021	TTCGGAGCCCACTTGTGAGTCAACAAAGAAATCTTCACTGTGA 1061	

DEFINITION	Mus musculus FEM, RhocGF A(hghe) and plackerin domain protein 1 (Chondrocyte-derived), mRNA (CDNA clone IMAGE:5374219), partial cds.
ACCESSION	BC043327
VERSION	BC043327.1
KEYWORDS	GI:28175181
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 2296) Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klauener, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Bluet, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heist, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Sleptom, M., Soares, M.B., Donald, M.F., Casavant, T.L., Scheet, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J., McEranon, K.D., Malek, J.A., Gunaratne, P.H., Richards, S., Woley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultk, S.W., Vittal, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Buettenfeld, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schermer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCES	2 (bases 1 to 2296)
AUTHORS	Straussberg, R.
JOURNAL	Direct Submission
TITLE	Submitted (09-JUN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNID) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smal, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNID at: <http://image.llnl.gov>
Series: IRAC Plate; 86 Row: 0 Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

RESULT 15			
BC043327			
LOCUS			
BC043327	2296 bp	mRNA	linear
			ROD 21-OCT-2003

```

FEATURES
source
Location/Qualifiers
1. .2296
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="Eva/N:10090"
/clone="IMAGE:5374219"
/tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month

```

old mouse. Taken by biopsy."
/clone.lib="NCI CGAP_Mam2"
/lab.host="DH10B"
/note="vector: PCMV-SPORT6"

gene
<1..2296
/gene="Farp1"
/note="synonym: Cdep"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
<1..926
/gene="Farp1"
/codon_start=3
/product="Farp1 protein"
/protein_id="AAH43327.1"
/db_xref="GI:28175182"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"

CDS
/translation="KLHELKDLIGIDNLVTRGRFIRLISLSKLGKLDQRMFPL
NDVLVTSKGLTASNQFKVHQLPLYQMTIESEBEMGVPHCLTRGQROSIIVASS
REMERKMDIOMALDLAEKSNPPELLASPPNDKSPDEATADQESDQLSART
SLERQAPHRGNTMVHVCWHRSTSVSVNDSIVENQSLRKRFKNSGMDKLWVF
TNPCLFYKSHODSHPLASLPILGYSLTPSESENIHKDYVKLHFKSHVYFRASE
VYFERMEVIRSATSSASRAHLHSKESHLV"
111..362
/gene="Farp1"
/note="PI; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PI domains have been found to possess inserted
domains (such as in PLC gamma, synctrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PI domain cause X-linked
agammaglobulinaemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CD:smart00233"
591..881
/gene="Farp1"
/note="PI; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PI domains have been found to possess inserted
domains (such as in PLC gamma, synctrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PI domain cause X-linked
agammaglobulinaemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"

misc_feature

591..881
/db_xref="CD:smart00233"
/gene="Farp1"
/note="PI; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PI domains have been found to possess inserted
domains (such as in PLC gamma, synctrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PI domain cause X-linked
agammaglobulinaemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CD:smart00233"

ORIGIN

Query Match 24.5%; Score 766.8; DB 6; Length 2296;
Best Local Similarity 89.5%; Pred. No. 2..6e-204;
Matches 825; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 2213 AGAAGCTGACGAACTCAAGAAAGATTGATGATGACATCTTGTGGTTCGGGAA 2272
DB 1 AGAAGCTGATGAGCTCAAGAAAGATCTGATCGGCATTGACATCTTGTGACCCGAGAA 60
QY 2273 GGAAGTTCAATCCGTCGGGCAAGCTCAGCAAGCTCAGGGAAGGGGCTCCAGACGCA 2332
DB 61 GGAAGTTCAATCCGTCGGGCAAGCTCAGCAAGCTCAGGGAAGGGGCTCCAGACGCA 120
QY 2333 TGTTCCTCTGTTCAACGACGTCCTCTATACAGAGCCGGGGCTGACGGCTCCATC 2392
DB 121 TGTTCCTCTGTTCAACGACGTCCTCTATACAGAGCCGGGGCTGACGGCTCCATC 180
QY 2393 AGTTTAAAGTCCAGGGGCAAGCTCCGCTCTATGCGATGACGATGAGAGAGGAAAGACG 2452
|||||

DB 181 AGTTTAAAGTCCAGGGGCAAGCTCCGCTCTATGCGATGACGATGAGAGAGTGAAGAG 240
QY 2453 AGTGGGGGGTGGCCCACTGCTGACCTCCGGGGGCGCAGCGAGTCCATCATGTTGGCCG 2512
DB 241 AGTGGGGGTGTGCCCACTGCTGACCTCCGGGGGCGCAGCGAGTCCATCATGTTGGCTG 300
QY 2513 CCAAGTTCTCGGTCGAGATGAGAAAGTGGGTTGAGAACATCCAGATGCGCATTTGACTGG 2572
DB 301 CCAAGTTCCCGGTCAAGATGAGAAAGTGGGTTGAGAACATCCAGATGCGCATTTGACTGG 360
QY 2573 CCGAAGAAAGAGCAGAGCCCGCCCGCTGAGTTCTCTGCGACGAGCCCGCTGACAAAGT 2632
DB 361 CAGAAAGAAAGCAATGAGACCCCGAGCTCTTGCCAGCAGCCCGCTGACAAAGT 420
QY 2633 CCCCTGATGAAGCCACCGCGGCTGACAGAGTCAAGAGATGACCTGAGCGGCTGCGCA 2692
DB 421 CCCCAGATGAAGCCACAGCGGCTGACAGAGTCAAGAGATGACCTGAGCTGCTCGGA 480
QY 2693 CATGCTGAGAGCGCAGGCGCCCGCAGCGCGCAACAAATGATGACGTGCTGGCACCC 2752
DB 481 CATGCTGAGAGCGCAGGCGCCCGCAGCGCGCAACAAATGATGACGTGCTGGCATC 540
QY 2753 GGAACACAGGCTCTCCATGATGAGACTTCAGATGCGAGTGAAGAAATCAAGTTGCGAA 2812
DB 541 GCAGACACAGGCTCTCCATGATGAGACTTCAGATGCGAGTGAAGAAATCAAGTTGCGAA 600
QY 2813 ACCCTGAGAGAAATTCAAAAACAGCAACGAGTGGCAGAGCTGAGTGTGTTACAA 2872
DB 601 ACTTCTGAGAAAGTTCAAGAACAGCAATGCGTGGCAGAGCTGAGTGTGTTACAA 660
QY 2873 ACTTCTGAGTGTGTTCTTCAAAATCAACACAGCAATGCGTGGCAGAGCTGAGTGTGTT 2932
DB 661 ACTTCTGAGTGTGTTCTTCAAAATCAACACAGCAATGCGTGGCAGAGCTGAGTGTGTT 720
QY 2933 TGCTGGGCTACTCCGCTCCACCAATCCCGCTGAGTCCGAGAACTCCAGAAAGACTAGTGT 2992
DB 721 TGCTGGGCTACTCCGCTCCACCAATCCCGCTGAGTCCGAGAACTCCAGAAAGACTAGTGT 780
QY 2993 TCAAGCTGCACTTCAAGTCCCAAGCTCTACTTCAAGGCGGAAAGCGAGTACAGCTTGC 3052
DB 781 TCAAGCTGCACTTCAAGTCCCAAGCTCTACTTCAAGGCGGAAAGCGAGTACAGCTTGC 840
QY 3053 AAAGGTGATGGAATGATCGCAGTGCACCAAGCTCTGCTCGGAGACCCACGTTGTTGA 3112
DB 841 AAAGGTGATGGAATGATCGCAGTGCACCAAGCTCTGCTCGGAGACCCACGATCTTGA 900
QY 3113 GCCACAAAGAGTCTTGTGTTA 3134
DB 901 GTCACAAAGAAATCTCACCCTGA 922

Search completed: May 22, 2006, 07:59:20
Job time : 11576.6 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2006, 00:29:52 ; Search time 1226.93 Seconds
(without alignments)
17815.260 Million cell updates/sec

Title: US-09-555-342b-1_COPY_49_3183

Perfect score: 3135
Sequence: 1 atcgaggaatagacagcagag.....acaaagagctctctgtgtat 3135

Scoring table: IDENTITY_NUC

Gapox 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: N_Geneseq_8: *
2: geneseqn1980s: *
3: geneseqn1980s: *
4: geneseqn2000s: *
5: geneseqn2001as: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *
14: geneseqn2005s: *
15: geneseqn2006s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3135	100.0	3442	2	AAX79183 Human cho
2	3135	100.0	3442	15	AEB03370 Acute mye
3	3133	99.9	5028	14	AEB22714 Human col
4	3131.8	99.9	4687	3	AAAO8582 Human cyt
5	3119.2	99.5	3187	5	AAS64760 DNA encod
6	1431.2	45.7	3094	3	AAC88992 Human pan
7	1037.4	33.1	3997	12	ADG84807 Human tum
8	1037.4	33.1	3997	13	ADG84807 Breast ca
9	1037.4	33.1	3997	13	ADG84807 Breast ca
10	550.4	17.6	582	10	ACD92278 Human col
11	550.4	17.6	582	10	ACD92278 Human col
12	447	14.3	3200	12	ADG63863 Novel hum
13	420.8	13.4	435	14	ACU61462 Human col
14	403.4	12.9	637	14	ACU61462 Human col
15	402.4	12.8	1718	4	AAS27016 Human enco
16	402.4	12.8	1718	10	ADG93194 Human CDN
17	385.6	12.3	393	9	ACH21582 Human adu
18	358.4	11.4	506	6	ABL67912 Ovary can

19	356.8	11.4	671	13	ADG52709 Novel can
20	334	10.7	441	15	AEB88734 Human CDN
21	306.6	9.8	3275	13	ADG07347 Full leng
22	304.8	9.7	2277	10	ADG71321 Human col
23	297	9.5	379	8	ABX34989 Bovine ES
24	292	9.3	400	8	ABX34991 Bovine ES
25	254	8.1	1824	5	AAS64761 DNA encod
26	233.4	7.4	6503	13	ACN41157 Human dia
27	229	7.3	1220	3	ABL02335 Drosophi1
28	223.6	7.1	241	4	AA42150 Human sec
29	207.8	6.6	716	4	AAS27440 Human enco
30	207.8	6.6	716	10	ADG93618 Human CDN
31	202.8	6.5	1747	6	ABL03867 Drosophi1
32	200.6	6.4	717	6	ABL00828 Human pol
33	192.4	6.1	2410	8	ACC44336 Gene enco
34	192.4	6.1	2636	14	AEA19659 Novel hum
35	192.4	6.1	2867	12	ADP10335 Reference
36	192.4	6.1	2867	13	ADR90365 Human ful
37	192.4	6.1	2867	13	ADP54688 Human PRO
38	192.4	6.1	2867	14	ADY15865 DNA encod
39	192.4	6.1	4338	14	AEA19657 Novel hum
40	192.4	6.1	4922	5	ABV30214 Human pro
41	187	6.0	2882	14	ADG2844 Murine pr
42	181.8	5.8	2781	14	ADY43235 Human psy
43	181.8	5.8	3643	6	ABK84074 Human CDN
44	181.8	5.8	3643	13	ADR40158 Human pro
45	181.8	5.8	3963	13	ADP54549 Human PRO

ALIGNMENTS

RESULT 1	
AAX79183	
ID	AAX79183 standard; DNA; 3442 BP.
XX	
XX	AAX79183;
AC	
XX	
DT	17-AUG-1999 (first entry)
XX	
DE	Human chondrocyte-derived gene CDEP.
XX	
KW	Differentiation; human; foetal; chondrocyte; ezrin-like domain; cancer;
KM	DBI homology domain; pleckstrin homology domain; rheumatoid arthritis;
XX	drug; ss.
OS	Homo sapiens.
PN	WO928458-A1.
XX	
PD	10-JUN-1999.
XX	
PF	27-NOV-1998; 98WO-JP005348.
XX	
PR	27-NOV-1997; 97JP-00342060.
XX	
PA	(CHUS) CHUGAI SEIYAKU KK.
XX	
PI	Kato Y, Kawamoto T, Koyano Y;
XX	
DR	WPI; 1999-371117/31.
XX	
DR	P-PSDB; AAY07482.
PT	Protein CDEP expressed in differentiated chondrocytes, and gene encoding
XX	it.
PS	Claim 5; Fig 1; 59pp; Japanese.
XX	
CC	This sequence represents the coding region for a protein (CDEP) expressed
CC	in differentiated human foetal chondrocytes, which contains an ezrin-like
CC	domain, a DBI homology (DH) domain and a pleckstrin homology (PH) domain.
CC	The nucleic acid or protein can be used in the investigation and
CC	treatment of cancers and arthritic diseases (including chronic rheumatoid

CC arthritis), or for screening of candidate anticancer drugs
XX Sequence 3442 BP; 864 A; 952 C; 927 G; 699 T; 0 U; 0 Other;

Query Match 100.0%; Score 3135; DB 2; Length 3442;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGGAATGATGAGAGAGGCGCCACCGAGATGACGACTGGGGGCCCCGGAAAAATTGG 60
DB |||||
QY 49 ATGGGGAATGATGAGAGAGGCGCCACCGAGATGACGACTGGGGGCCCCGGAAAAATTGG 108
DB |||||
QY 61 GGGATCAGTACCTTGGAGCGTGGACAGAAAGCCGCCCAACACCTTCAGAAAAATCTGTG 120
DB |||||
QY 109 GGGATCAGTACCTTGGAGCGTGGACAGAAAGCCGCCCAACACCTTCAGAAAAATCTGTG 168
DB |||||
QY 121 TCCATCAAAATCCAGATGCTGATGACCCGAGAGGCACTTTGAAGTTCCACAAAAGAGT 180
DB |||||
QY 169 TCCATCAAAATCCAGATGCTGATGACCCGAGAGGCACTTTGAAGTTCCACAAAAGAGT 228
DB |||||
QY 181 CCTGGAGAGTGTGCTGATGATGACCTTCACCTCACTCGTGGAAAGTGAAT 240
DB |||||
QY 229 CTTGGAGAGTGTGCTGATGATGACCTTCACCTCGTGGAAAGTGAAT 288
DB |||||
QY 241 TTTGGCTCGAGTTTCTGATCACAAAAAGATCAGGTGTGCTGATCTCTAAAACC 300
DB |||||
QY 289 TTTGGCTCGAGTTTCTGATCACAAAAAGATCAGGTGTGCTGATCTCTAAAACC 348
DB |||||
QY 301 ATTGTAAACAGATTGAAAGGCCAAGCACTGTGTGTTAAGTTGTGTGAAATCTTT 360
DB |||||
QY 349 ATTGTAAACAGATTGAAAGGCCAAGCACTGTGTGTTAAGTTGTGTGAAATCTTT 408
DB |||||
QY 361 CGGCTGACACACACAACTCCAAAGAAAGTCAACAAGGTAACCTGTTCGCGTGCAGGTG 420
DB |||||
QY 409 CGGCTGACACACACAACTCCAAAGAAAGTCAACAAGGTAACCTGTTCGCGTGCAGGTG 468
DB |||||
QY 421 AAGCAGAGCTGGCTCAAGGCAAGTTGACGTGTATGACACAGCGCAGCTCTCTGATT 480
DB |||||
QY 469 AAGCAGAGCTGGCTCAAGGCAAGTTGACGTGTATGACACAGCGCAGCTCTCTGATT 528
DB |||||
QY 481 TCACACATTGTGCAATCTGAGATTGGGGATTTTGAAGCCTTGGACAGAGCACTTA 540
DB |||||
QY 529 TCACACATTGTGCAATCTGAGATTGGGGATTTTGAAGCCTTGGACAGAGCACTTA 588
DB |||||
QY 541 GCAAAAATTAATATACCTCTCAGAAAGCAGCTAGAGGACAAATTCGTGAAATTTTCA 600
DB |||||
QY 589 GCAAAAATTAATATACCTCTCAGAAAGCAGCTAGAGGACAAATTCGTGAAATTTTCA 648
DB |||||
QY 601 CATTAACCATTTGACAAACACCAAGAGATCAGATTTTCAAGTCTTAGAGATTGCCGT 660
DB |||||
QY 649 CATTAACCATTTGACAAACACCAAGAGATCAGATTTTCAAGTCTTAGAGATTGCCGT 708
DB |||||
QY 661 CGGCTAGAGATGTATGGAATCCGGTTGCAACCGGCCAAGGACAGGAAAGGCAAGAGATC 720
DB |||||
QY 709 CGGCTAGAGATGTATGGAATCCGGTTGCAACCGGCCAAGGACAGGAAAGGCAAGAGATC 768
DB |||||
QY 721 AATCTGCGCTGGCCAAACAGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
DB |||||
QY 769 AATCTGCGCTGGCCAAACAGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 828
DB |||||
QY 781 TTCAACTGGGCCCAAGGTGCGGAAAGCTGAGCTTCAAGAGAAAGCGCTTTCATCAAGCTC 840
DB |||||
QY 841 CGGCGCAATGCGCAATAGTCCGTAACAGAGTACCTTGGAAATTCCTGATGCGCAGTCCGGAT 900
DB |||||
QY 889 CGGCGCAATGCGCAATAGTCCGTAACAGAGTACCTTGGAAATTCCTGATGCGCAGTCCGGAT 948
DB |||||
QY 901 TTCTGCAAGTCTTTCGAAAAATCTGTGTGAACATCATGCTTCTTGAATCTTTTGA 960
DB |||||
QY 949 TTCTGCAAGTCTTTCGAAAAATCTGTGTGAACATCATGCTTCTTGAATCTTTTGA 1008
DB |||||
QY 961 GAGCCCAAAACAAAGCCCAAGCCGTCCTTTAGCCGGGGTCAATCATTTGCTGAGT 1020
DB |||||
```

```
DB |||||
QY 1009 GAGCCCAAAACAAAGCCCAAGCCGTCCTTTACCGGGGTCATCATTTGCTGATCAGT 1068
DB |||||
QY 1021 GGTCCGACCTCAGAAACAGAGTTTTCAGATATGTTAAAGAGAGACATTAAGAAAGTGCAG 1080
DB |||||
QY 1069 GGTCCGACCTCAGAAACAGAGTTTTCAGATATGTTAAAGAGAGACATTAAGAAAGTGCAG 1128
DB |||||
QY 1081 TTTGAAAGGAGCAGACAGCAAGATTCATTTCTATCCGAGCCTTGCTTCACAGCCTACAGA 1140
DB |||||
QY 1129 TTTGAAAGGAGCAGACAGCAAGATTCATTTCTATCCGAGCCTTGCTTCACAGCCTACAGA 1188
DB |||||
QY 1141 CTGAATTCGAAAGTGTGAGAGCTCTCAGACAGACCAAGCTTACATTTGAGAGAGT 1200
DB |||||
QY 1189 CTGAATTCGAAAGTGTGAGAGCTCTCAGACAGACCAAGCCTTACATTTGAGAGAGT 1248
DB |||||
QY 1201 GCGGAATCTCCAGGGGGCCAGAGCTGCGGGCCAGAGAAAGGAAAGGTTTCCGCGGG 1260
DB |||||
QY 1249 GCGGAATCTCCAGGGGGCCAGAGCTGCGGGCCAGAGAAAGGAAAGGTTTCCGCGGG 1308
DB |||||
QY 1261 GAGCCGGGGTCCACACCCGAGCCCTGCGCGAGAGAAAGCCCGGGTTAAACAAGAGGCG 1320
DB |||||
QY 1309 GAGCCGGGGTCCACACCCGAGCCCTGCGCGAGAGAAAGCCCGGGTTAAACAAGAGGCG 1368
DB |||||
QY 1321 GACGAGCCGCTCGGCGCCACGAGAGAAAGAGAGAGTGTTAAGATAGACCCAG 1380
DB |||||
QY 1369 GACGAGCCGCTCGGCGCCACGAGAGAAAGAGAGTGTTAAGATAGACCCAG 1428
DB |||||
QY 1381 CAGAGTAAACCTCACGCCCCGAGCCAGACAGAGCTCCTGACTGGCAGTCTCACTT 1440
DB |||||
QY 1429 CAGAGTAAACCTCACGCCCCGAGCCAGACAGAGCTCCTGACTGGCAGTCTCACTT 1488
DB |||||
QY 1441 TCCGAGCTGTCTGTAACTCGAGGGGGAGTGTGACCTCTCCAAAGTGAACCTGTCTCC 1500
DB |||||
QY 1489 TCCGAGCTGTCTGTAACTCGAGGGGGAGTGTGACCTCTCCAAAGTGAACCTGTCTCC 1548
DB |||||
QY 1501 AACCTGAGCCCGGACACCAAGAGGCTCTCCCTTGAATCAGCCGCTGTGAATGACAG 1560
DB |||||
QY 1549 AACCTGAGCCCGGACACCAAGAGGCTCTCCCTTGAATCAGCCGCTGTGAATGACAG 1608
DB |||||
QY 1561 GCTGCGCCCGGAGCGAGATGAGAGTGAAGGCGGAGGAAAGATTCCCACTGATAA 1620
DB |||||
QY 1609 GCTGCGCCCGGAGCGAGATGAGAGTGAAGGCGGAGGAAAGATTCCCACTGATAA 1668
DB |||||
QY 1621 GCGTACTTATAGCTAAGAGAGTGTCTACACAGAGCAATATCTGAAGATCTCGAA 1680
DB |||||
QY 1669 GCGTACTTATAGCTAAGAGAGTGTCTACACAGAGCAATATCTGAAGATCTCGAA 1728
DB |||||
QY 1681 GTTATCATCTTGTGTTTCAAGCAAGTGAAGAAAGAGAGCCATGCGGAAAGCATCTG 1740
DB |||||
QY 1729 GTTATCATCTTGTGTTTCAAGCAAGTGAAGAAAGAGAGCCATGCGGAAAGCATCTG 1788
DB |||||
QY 1741 AAAAGTCTCATATTCGCCGAATTTTGAACCTTGGCAAAATTTCTACTAATTTTCTCAAG 1800
DB |||||
QY 1789 AAAAGTCTCATATTCGCCGAATTTTGAACCTTGGCAAAATTTCTACTAATTTTCTCAAG 1848
DB |||||
QY 1801 GAAATTTGAGCAACGACTTGCCTGTGAGAAAGGCGCTCAATATGCCAAATTCAGAGTTAC 1860
DB |||||
QY 1849 GAAATTTGAGCAACGACTTGCCTGTGAGAAAGGCGCTCAATATGCCAAATTCAGAGTTAC 1908
DB |||||
QY 1861 CAAAGATTCGCGAGTGTCTAGTGAAGAACTTCAGGGCATTAAGACCTTGGCGGTCAAC 1920
DB |||||
QY 1909 CAAAGATTCGCGAGTGTCTAGTGAAGAACTTCAGGGCATTAAGACCTTGGCGGTCAAC 1968
DB |||||
QY 1921 CTGTGGAAGCAGGAGGCTTTGAGAGCCCTTGAGAAATGAAATCAAGAGCTCCGGCGG 1980
DB |||||
QY 1969 CTGTGGAAGCAGGAGGCTTTGAGAGCCCTTGAGAAATGAAATCAAGAGCTCCGGCGG 2028
DB |||||
QY 1981 CTGAGAACTTCTGCAAGACTTTGAGCTGCAAGAGTGTGTACTTACCTCAACCTCAACAC 2040
DB |||||
QY 2029 CTGAGAACTTCTGCAAGACTTTGAGCTGCAAGAGTGTGTACTTACCTCAACCTCAACAC 2088
DB |||||
QY 2041 TTCTCTCGGGCCACTGACCGGCTCATGCACTCAAGCAGGTCTTGAAGGCGCTGTGC 2100
DB |||||
```



```

Db 2089 TTCTCTGCGGCACTGACACCGGCTCATGACTACAAAGAGGTCCTGAGCGGCTGTGC 2148
Qy 2101 AAACACCAACCCCGGAGCCACGCCGACTTCAGGGAGCTGCCGAGCCGTTTGGCAGAGATC 2160
Db 2149 AAACACCAACCCCGGAGCCACGCCGACTTCAGGGAGCTGCCGAGCCGTTTGGCAGAGATC 2208
Qy 2161 ACCGAGATGCTGACAGCTCCACGGATGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2220
Db 2209 ACCGAGATGCTGACAGCTCCACGGATGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2268
Qy 2221 CACGAATCAAGAAAGATTTGATTGGCATTGACATCTTGTGTTCCGGGAGAGAGTTTC 2280
Db 2269 CACGAATCAAGAAAGATTTGATTGGCATTGACATCTTGTGTTCCGGGAGAGAGTTTC 2328
Qy 2281 ATCCGTTGCGGAGCCCTCAGCAAGCTTCGGGGAAAGGGGCTCCAGCAGCCGATGTTCTTC 2340
Db 2329 ATCCGTTGCGGAGCCCTCAGCAAGCTTCGGGGAAAGGGGCTCCAGCAGCCGATGTTCTTC 2388
Qy 2341 CTGTTCAACGACGTCCTGCTATACACGAGCCGGGGGCTGACGGGCTTCATCAGTTTAA 2400
Db 2389 CTGTTCAACGACGTCCTGCTATACACGAGCCGGGGGCTGACGGGCTTCATCAGTTTAA 2448
Qy 2401 GTTCACGGGAGGTCCTCCGCTCTATGGCATGACGATTGAGAGAGCGAAGACGAGTGGGG 2460
Db 2449 GTTCACGGGAGGTCCTCCGCTCTATGGCATGACGATTGAGAGAGCGAAGACGAGTGGGG 2508
Qy 2461 GTGCCCCACTGCTGACCTTCGCGGGCCAGCGGCACTCATCTCTGTGGCCGCAAGTTCT 2520
Db 2509 GTGCCCCACTGCTGACCTTCGCGGGCCAGCGGCACTCATCTCTGTGGCCGCAAGTTCT 2568
Qy 2521 CGGTCCGAGATGAGAAAGTGGGTTGAGGACATCGAGTGGCCATTGACCTGCGGAGAG 2580
Db 2569 CGGTCCGAGATGAGAAAGTGGGTTGAGGACATCGAGTGGCCATTGACCTGCGGAGAG 2628
Qy 2581 AGCAGACGCCCGCCCTGAGTTCTGTGGCAGAGCCCGCTGACAAAGTCCCTGAT 2640
Db 2629 AGCAGACGCCCGCCCTGAGTTCTGTGGCAGAGCCCGCTGACAAAGTCCCTGAT 2688
Qy 2641 GAAGCCACCGGCTGACCGAGATGAGAGATGACTGAGCCCTTCGCGCAATGGCTG 2700
Db 2689 GAAGCCACCGGCTGACCGAGATGAGAGATGACTGAGCCCTTCGCGCAATGGCTG 2748
Qy 2701 GAGCGCAGGCGCCCGCAGCGCGGCAACAATGATGACGTGTGCTGGACCGCAACCC 2760
Db 2749 GAGCGCAGGCGCCCGCAGCGCGGCAACAATGATGACGTGTGCTGGACCGCAACCC 2808
Qy 2761 AGCGTCTCCATGATGATCTTCAAGATCGAGTGAAGATCAGTTGTCTGAAACCTGCTG 2820
Db 2809 AGCGTCTCCATGATGATCTTCAAGATCGAGTGAAGATCAGTTGTCTGAAACCTGCTG 2868
Qy 2821 AGGAATTTCAAAAACGCAACGGGTGGCAGAGCTGTGGTGTGTTCAAACTTCTGC 2880
Db 2869 AGGAATTTCAAAAACGCAACGGGTGGCAGAGCTGTGGTGTGTTCAAACTTCTGC 2928
Qy 2881 CTGTTCTTCAAAATCACACGAGCAATCATCCCTTSCAGGCTGCTGCTCGG 2940
Db 2929 CTGTTCTTCAAAATCACACGAGCAATCATCCCTTSCAGGCTGCTGCTCGG 2988
Qy 2941 TACTGCTCAACATCCCTCTGATCGGAAACATCCAGAAAGATCAAGTTTCAAGCTG 3000
Db 2989 TACTGCTCAACATCCCTCTGATCGGAAACATCCAGAAAGATCAAGTTTCAAGCTG 3048
Qy 3001 CACTTCAAGTCCCACTCTACTACTTCAAGGCGGAAAGGAGTACACGTTTCAAAAGCTG 3060
Db 3049 CACTTCAAGTCCCACTCTACTACTTCAAGGCGGAAAGGAGTACACGTTTCAAAAGCTG 3108
Qy 3061 ATGGAAGTATCCGAGTGCACCAAGCTTGCCTCGGAAACCCCAAGTGTGAGCCACAA 3120
Db 3109 ATGGAAGTATCCGAGTGCACCAAGCTTGCCTCGGAAACCCCAAGTGTGAGCCACAA 3168
Qy 3121 GAGTCTCTGTGTAT 3135
Db 3169 GAGTCTCTGTGTAT 3183

```

```

RESULT 2
AEB03370
ID AEB03370 standard; DNA; 3442 BP.
XX
AC AEB03370;
XX
DT 09-FEB-2006 (first entry)
XX
DE Acute myelogenous leukemia prognosis related DNA sequence SEQ ID NO: 31.
XX
KW ds: gene; acute myelogenous leukemia; prognosis; gene expression;
KW bioclip.
XX
OS Homo sapiens.
XX
PN JF200533987-A.
XX
PD 08-DEC-2005.
XX
PF 06-MAY-2005; 2005JP-00135284.
XX
PR 06-MAY-2004; 2004US-0568635P.
XX
PA (VERI-) VERIDEX LLC.
XX
PI Laponi M.
XX
DR WPI; 2006-004067/01.
XX
PT Evaluating an acute myelogenous leukemia patient's prognosis, comprising
PT detecting lower or higher expression level of gene recognized by probe
PT set e.g. 202820-at and 206148-at, with respect to predetermined cut-off
PT level.
XX
PS Example 5; SEQ ID NO 31; 60pp; Japanese.
XX
XX
CC The present sequence is that of a human gene sequence which is claimed
CC for use in evaluating the prognosis of patients suffering from acute
CC myelogenous leukemia (AML) by analysis of the patients gene expression
CC profile. The invention relates to a novel method for evaluating the
CC prognosis of a patient with acute myelogenous leukemia by detecting
CC higher/lower expression level of genes encoding mRNA recognized by a
CC probe set chosen from 19 nucleotide sequences (AEB03340 or AEB03344-
CC AEB03361). Also, gene expression profiling of AML patients using a set of
CC probes based on 167 sequences (AEB03362-AEB03527) using a bioclip was
CC performed to determine prognosis. Methods are also included for
CC determining an AML patient's state, determining AML patient's treatment,
CC protocol, determining whether the patient will respond to the treatment,
CC and producing an AML patient's prognosis report by analysis of the
CC patients gene expression profile. The methods are useful in evaluating an
CC AML patient's prognosis and for treating AML patient.
XX
SQ Sequence 3442 BP; 864 A; 952 C; 927 G; 699 T; 0 U; 0 Other;
Qy Query Match 100.0%; Score 3135; DB 15; Length 3442;
Db Best Local Similarity 100.0%; Pred. No. 0;
Matches 3135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGGAATATGAGCAGAGCGGCAACCCAGGATCAAGCTGGGGCCCCGGAAATTCG 60
Db 49 ATGGGGAATATGAGCAGAGCGGCAACCCAGGATCAAGCTGGGGCCCCGGAAATTCG 108
Qy 61 GGGATCAGTACTTGGAAAGTGAACAAGCCGCCCAACCTTCAGAAAACTCGTG 120
Db 109 GGGATCAGTACTTGGAAAGTGAACAAGCCGCCCAACCTTCAGAAAACTCGTG 168
Qy 121 TCCATCAAAATTCAGATGTGATGACACCCAGAGGCAATTTGAAGTTCCAAAGAGCT 180
Db 169 TCCATCAAAATTCAGATGTGATGACACCCAGAGGCAATTTGAAGTTCCAAAGAGCT 228
Qy 181 CTTGGGAAGTGTGCTGATGCAATTGGCAACCACTCAACCTGTGGAAGTGAATAT 240

```


229 CCGGGAAGTGTGTGATGACGATTGCAACCACTCAACCTCGTGAAGGTGACTAT 288
241 TTTGGCTCGAGTTTCTCTGATCACAAAAATTCACGGTGTGGTGGATCTCTAAAACCC 300
289 TTTGGCTCGAGTTTCTCTGATCACAAAAATTCACGGTGTGGTGGATCTCTAAAACCC 348
301 ATTGTGAACAGATTGAAGGCCAAGCAAGTGTGTTAAGTTTGGTGAATTTCTT 360
349 ATTGTGAACAGATTGAAGGCCAAGCAAGTGTGTTAAGTTTGGTGAATTTCTT 408
361 CCGCTGACCAACACCACTCCAAAGAACTCACAAAGTACCTGTTCCGCTGACAGTG 420
409 CCGCTGACCAACACCACTCCAAAGAACTCACAAAGTACCTGTTCCGCTGACAGTG 468
421 AAGCAGAGATTGGCTCAAGGAGGTGAGGTGAAGACCAAGCCGACCTCTTGAT 480
469 AAGCAGAGATTGGCTCAAGGAGGTGAGGTGAAGACCAAGCCGACCTCTTGAT 528
481 TCACACATTGTGCAATCTGAGATTGGGGAATTTGATGAAGCTTGGACAGAGCACTTA 540
529 TCACACATTGTGCAATCTGAGATTGGGGAATTTGATGAAGCTTGGACAGAGCACTTA 588
541 GCAAAAATTAATATACATACCTCAGCAAGACGCACTAGAGACAAAATCGTGAATTTGAC 600
589 GCAAAAATTAATATACATACCTCAGCAAGACGCACTAGAGACAAAATCGTGAATTTGAC 648
601 CATTAACCAATTGGACAAACACAGAGAAATGCAATTTCCAGTCTCTAGAAATGGCCGT 660
649 CATTAACCAATTGGACAAACACAGAGAAATGCAATTTCCAGTCTCTAGAAATGGCCGT 708
661 CCGCTAGAGATGATGAATCCGGTGGCAACCCGGCCAAAGACAGAGAGCAAGAGATC 720
709 CCGCTAGAGATGATGAATCCGGTGGCAACCCGGCCAAAGACAGAGAGCAAGAGATC 768
721 AATCTGGCCGTGGCCAAACAGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGTC 780
769 AATCTGGCCGTGGCCAAACAGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGTC 828
781 TTCAACTGGGCCAAGGTGGGAAAGCTGAGCTTCAAGAGAAAGCTTTCATCAAGCTC 840
829 TTCAACTGGGCCAAGGTGGGAAAGCTGAGCTTCAAGAGAAAGCTTTCATCAAGCTC 888
841 CCGCCAGATGCCAATAGTGCATACAGAGATACCTTGAATTCCTGATGGCCAGTCCGGAT 900
889 CCGCCAGATGCCAATAGTGCATACAGAGATACCTTGAATTCCTGATGGCCAGTCCGGAT 948
901 TTCTGCAAGTCTTCTGGAATAATCTGTGTTGAACATCATGCTTCTTTAGACTTTTGA 960
949 TTCTGCAAGTCTTCTGGAATAATCTGTGTTGAACATCATGCTTCTTTAGACTTTTGA 1008
961 GAGCCCAAAACCAAGGCCAAGCCGTCCTTTAGCCGGGGGTCAATTTGGGTTCACT 1020
1009 GAGCCCAAAACCAAGGCCAAGCCGTCCTTTAGCCGGGGGTCAATTTGGGTTCACT 1068
1021 GGTTCGCACTCAGAGAGGTTCTGACTATGTGTTAAAGAGAGACATTAAGAGGTGAC 1080
1069 GGTTCGCACTCAGAGAGGTTCTGACTATGTGTTAAAGAGAGACATTAAGAGGTGAC 1128
1081 TTTGAAAGAGAGACAGCAAGATTCAATTCATCCGAGCCCTTGCTTCAAGCTTACAGAA 1140
1129 TTTGAAAGAGAGACAGCAAGATTCAATTCATCCGAGCCCTTGCTTCAAGCTTACAGAA 1188
1141 CTGAATTCGGAAGTGTGAGAGCTCAGCAGAGACCAAGCTTCAATTTGGAGAAAGGT 1200
1189 CTGAATTCGGAAGTGTGAGAGCTCAGCAGAGACCAAGCTTCAATTTGGAGAAAGGT 1248
1201 GCGCAATCTTCAGAGGGGCGAGAGCTCGCGCAGAGAAAGAAACGAAAGGTTTCCGCGGG 1260
1249 GCGCAATCTTCAGAGGGGCGAGAGCTCGCGCAGAGAAAGAAACGAAAGGTTTCCGCGGG 1308
1261 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTTAACAGCAGGGG 1320

1309 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTTAACAGCAGGGG 1368
1321 GAGCGAGCGGCTCGGCGGCCACGAGAGAAAGAGAGAGGTGTTAAGATAGAGCCCG 1380
1369 GAGCGAGCGGCTCGGCGGCCACGAGAGAAAGAGAGAGGTGTTAAGATAGAGCCCG 1428
1381 CAGAGTAAACCTCAGCCCCCGAGCCAAAGACAGAGCTCCCTGACTGGCAGTCTCACTT 1440
1429 CAGAGTAAACCTCAGCCCCCGAGCCAAAGACAGAGCTCCCTGACTGGCAGTCTCACTT 1488
1441 TCCGAGCTGTCTGTGAATCTCGAGGGGGAGTGGCCCCCTGCCAACGTACTTGTCTCC 1500
1489 TCCGAGCTGTCTGTGAATCTCGAGGGGGAGTGGCCCCCTGCCAACGTACTTGTCTCC 1548
1501 AACCTGAGCCCGACACCAAGAGGCTCTCCCTTGATCAGCCGCTGTAATGACAG 1560
1549 AACCTGAGCCCGACACCAAGAGGCTCTCCCTTGATCAGCCGCTGTAATGACAG 1608
1561 GCTTCCCGGACGAGCATGAGAGATGAGGGCCGAGAGAAAGATTCCCACTGATAAA 1620
1609 GCTTCCCGGACGAGCATGAGAGATGAGGGCCGAGAGAAAGATTCCCACTGATAAA 1668
1621 GGTACTTCTATAGCTTAAGAAAGTCTTACACCGAGCGCAATATCTGAAGATCTCGAA 1680
1669 GGTACTTCTATAGCTTAAGAAAGTCTTACACCGAGCGCAATATCTGAAGATCTCGAA 1728
1681 GTTATCACTTGTGTGTTTCAAGCACAGTGAAGCAAAAGAGAGCGCATGCCGAGACACTG 1740
1729 GTTATCACTTGTGTGTTTCAAGCACAGTGAAGCAAAAGAGAGCGCATGCCGAGACACTG 1788
1741 AAAAGTCTCATATTTCCGAATTTTGAACCTTTGCAAAATTTCACTAATTTTCTCAAG 1800
1789 AAAAGTCTCATATTTCCGAATTTTGAACCTTTGCAAAATTTCACTAATTTTCTCAAG 1848
1801 GAAATTTGAGCAACAATTGGCCCTGTGGGAAGGCGCTCAATGCCCAATCAAGATTAAC 1860
1849 GAAATTTGAGCAACAATTGGCCCTGTGGGAAGGCGCTCAATGCCCAATCAAGATTAAC 1908
1861 CAAGAATTCGGGAGTGTATGCTGTAAGAAATTCAGGGCAATGAACACCTGGCGGCTAC 1920
1909 CAAGAATTCGGGAGTGTATGCTGTAAGAAATTCAGGGCAATGAACACCTGGCGGCTAC 1968
1921 CTGTGAAGCAACAGGAGGCTTTGAGGCGCTGAGAAATGAAATCAAGACTCCGGCGG 1980
1969 CTGTGAAGCAACAGGAGGCTTTGAGGCGCTGAGAAATGAAATCAAGACTCCGGCGG 2028
1981 CTGTGAAGCACTTGTGAAGACTTTGAGGCTGCAAGAAAGGTGTATCTACCGCTCAACAC 2040
2029 CTGTGAAGCACTTGTGAAGACTTTGAGGCTGCAAGAAAGGTGTATCTACCGCTCAACAC 2088
2041 TTCTCTGCGGCACTGCAACCGGCTCATGCACTAACAGAGGTCTGAGAGCGGCTGTGC 2100
2089 TTCTCTGCGGCACTGCAACCGGCTCATGCACTAACAGAGGTCTGAGAGCGGCTGTGC 2148
2101 AAAACCAACCCGCGCAGGCAAGCTTCAAGGACTGCGGCTTGGCAGAGATC 2160
2149 AAAACCAACCCGCGCAGGCAAGCTTCAAGGACTGCGGCTTGGCAGAGATC 2208
2161 ACGAGAGATGGGCAACACTCCAGGTACAGATGAATCAAGATGAGAAATTTCCAGAAAGCTG 2220
2209 ACGAGAGATGGGCAACACTCCAGGTACAGATGAATCAAGATGAGAAATTTCCAGAAAGCTG 2268
2221 CACGAATCAAGAAAGATTGATTTGGATTGACAAATCTTGTGTGTTCCGGAGAGGAGTTT 2280
2269 CACGAATCAAGAAAGATTGATTTGGATTGACAAATCTTGTGTGTTCCGGAGAGGAGTTT 2328
2281 ATCCGTCTGGGCAAGCTTCAAGCAAGCTTCTCGGGAAAGGAGCTTCACAGCGCATGTTCTT 2340
2329 ATCCGTCTGGGCAAGCTTCAAGCAAGCTTCTCGGGAAAGGAGCTTCACAGCGCATGTTCTT 2388
2341 CTGTTCAACGACGTCTCTGTTATACAGAGCCGGGGGCTGACGGCTCCAAATCAATTTAAA 2400
2389 CTGTTCAACGACGTCTCTGTTATACAGAGCCGGGGGCTGACGGCTCCAAATCAATTTAAA 2448

QY	2401	GTCCACGGGGACGCTCCCGCTCTATAGGACATGACGATTTGAGAGAGCGAAGACGATGGGGG	2466
Db	2449	GTCCACGGGGACGCTCCCGCTCTATAGGACATGACGATTTGAGAGAGCGAAGACGATGGGGG	2508
QY	2461	GTGCCCCCACTGCTCCGACCCCTCCGGGGCCAGAGCGGACAGTCCATCATTCGTGGCCGCGCATCTT	2520
Db	2509	GTGCCCCCACTGCTCCGACCCCTCCGGGGCCAGCGGACAGTCCATCATTCGTGGCCGCGCATCTT	2568
QY	2521	CGGTTCCGAGATGAGAAAGTGGGTGGAGGACATCCAGATGCGCCATTGACCTTGGCGGAGAG	2580
Db	2569	CGGTTCCGAGATGAGAAAGTGGGTGGAGGACATCCAGATGCGCCATTGACCTTGGCGGAGAG	2628
QY	2581	AGCAGCAGCCCCCGCCCTGAGATTCTTGCGCCAGACGCCCCCTTGACAAACAAATCCCTCTGAT	2640
Db	2629	AGCAGCAGCCCCCGCCCTGAGATTCTTGCGCCAGACGCCCCCTTGACAAACAAATCCCTCTGAT	2688
QY	2641	GAAGCCACCCGCGGCTGACAGGAGTCAAGAGATGACTGAGCGGCTCCGCGACATTCGCTG	2700
Db	2689	GAAGCCACCCGCGGCTGACAGGAGTCAAGAGATGACTGAGCGGCTCCGCGACATTCGCTG	2748
QY	2701	GAGCGCCAGGCCCGCGACCGCGGCAACAAATGTTGCACTGTGTCTTGGACCCGCAACCC	2760
Db	2749	GAGCGCCAGGCCCGCGACCGCGGCAACAAATGTTGCACTGTGTCTTGGACCCGCAACCC	2808
QY	2761	AGCGTCTTCATGTGTGAGACTTACGACATCCGACATGAGAAATCACTGTTCTTGGAAACTCTGTTG	2820
Db	2809	AGCGTCTTCATGTGTGAGACTTACGACATCCGACATGAGAAATCACTGTTCTTGGAAACTCTGTTG	2868
QY	2821	AGGAAATTCAAAAAACAGCAACGGGTGGGAGAAAGCTGTGGGTGGTGTTCACAAACTTCTGC	2880
Db	2869	AGGAAATTCAAAAAACAGCAACGGGTGGGAGAAAGCTGTGGGTGGTGTTCACAAACTTCTGC	2928
QY	2881	CTGTGTTCTTCAAAATCAACACAGGACAAATCATCCCTTGGCAAGCTTGCCTCTCGGCG	2940
Db	2929	CTGTGTTCTTCAAAATCAACACAGGACAAATCATCCCTTGGCAAGCTTGCCTCTCGGCG	2988
QY	2941	TACTCTGCTCAGCAATCCCTCTGAGATGCCAGAAACATCCGAAAGACTAGTGTTCAGAGTG	3000
Db	2989	TACTCTGCTCAGCAATCCCTCTGAGATGCCAGAAACATCCGAAAGACTAGTGTTCAGAGTG	3048
QY	3001	CACCTTCAAGTCCCAAGTCTACTCTTCAAGGCGGAAACGAGATCAAGTTCGAAAGGTGG	3060
Db	3049	CACCTTCAAGTCCCAAGTCTACTCTTCAAGGCGGAAACGAGATCAAGTTCGAAAGGTGG	3108
QY	3061	ATGGAAGTGAATCCGCAAGTCCACCAAGCTCTGCGCAGACCCCAAGTGTGAGCCACAA	3120
Db	3109	ATGGAAGTGAATCCGCAAGTCCACCAAGCTCTGCGCAGACCCCAAGTGTGAGCCACAA	3168
QY	3121	GAGTCTCTGTGTAT 3135	
Db	3169	GAGTCTCTGTGTAT 3183	

RESULT 3	
AEB22714	
ID	AEB22714 standard; DNA; 5028 BP.
XX	
AC	AEB22714;
XX	
DT	22-SEP-2005 (first entry)
XX	
DE	Human colon cancer-associated gene SEQ ID NO:9.
XX	
KW	cancer; genetic marker; lymphoma; cytostatic; hematological disease;
KW	immune disorder; neoplasm; colon tumor; gastrointestinal disease;
KW	adenocarcinoma; carcinoma; screening; diagnostic; prognostic;
KW	pharmaceutical; ds; gene.
XX	
OS	Homo sapiens.
XX	
PN	WO2005064009-A1.
XX	

PD		14-JUL-2005.
PP		
PR	23-DEC-2004;	2004WO-DK000914.
XX		
XX	27-DEC-2003;	2003DK-00001940.
PR	24-JAN-2004;	2004DK-00000096.
PR	07-APR-2004;	2004DK-00000586.
PR	26-NOV-2004;	2004DK-00001843.
XX		
PA	(AROS-) AROS APPLIED BIOTECHNOLOGY APS.	
XX		
P1	Ornloft TF, Jensen JL, Krühoffer M, Iahio P, Aaltonen LA;	
XX		
DR	WPI, 2005-506661/51.	
XX		
PT	Classifying the cancer in an individual having contracted cancer	
FP	comprises determining the microsatellite status of the tumor.	
PS	Claim 9; SEQ ID NO 9; 338pp; English.	
XX		
CC	The invention relates to a method of classifying the cancer in an	
CC	individual having contracted cancer, which comprises determining the	
CC	microsatellite status of the tumor. The microsatellite status or the	
CC	hereditary or sporadic nature of the cancer is determined by a method	
CC	comprising: (a) obtaining a sample from the individual having contracted	
CC	cancer; the sample comprising gene expression products, the presence	
CC	and/or amount of which forms a pattern that is indicative of the	
CC	microsatellite status or the hereditary or sporadic nature of the cancer;	
CC	(b) determining the presence and/or amount of the gene expression	
CC	products forming the pattern, and obtaining an indication of the	
CC	microsatellite status or the hereditary or sporadic nature of the cancer	
CC	in the individual based on step (b). At least one or two of the gene	
CC	expression products forming a pattern used to determine the	
CC	microsatellite status or the hereditary or sporadic nature of the cancer	
CC	are selected individually from any of the 134 genes comprising 367-7850	
CC	bp (ABE22706 to ABE2839), given in the specification. Also included are:	
CC	a method for treating an individual; a method for reducing malignancy of	
CC	a cell; a method for producing antibodies against an expression product	
CC	of a cell from a biological tissue; pharmaceutical composition, for	
CC	treating a classified cancer, comprising at least one antibody, at least	
CC	one polypeptide, or at least one nucleic acid and/or probe; and an assay,	
CC	for classification of cancer in an individual having contracted cancer.	
CC	comprising at least one marker capable of determining the microsatellite	
CC	status in a sample and at least one marker in a sample determining the	
CC	prognostic marker, where the microsatellite status and the prognostic	
CC	marker is determined simultaneously or sequentially. The method above is	
CC	useful for producing an assay for classifying cancer in animal tissue.	
CC	The peptide, gene, or probe is useful for preparing a pharmaceutical	
CC	composition for the treatment of a cancer in animal tissue. The method is	
CC	useful for classifying the cancer in an individual having contracted	
CC	cancer, i.e. colon cancer, e.g. an adenocarcinoma, a carcinoma, a	
CC	teratoma, a sarcoma, and/or a lymphoma. The present sequence represents a	
CC	colon cancer-associated gene used in the method of the invention.	
XX		
SO	Sequence 5028 BP; 1288 A; 1387 C; 1276 G; 1077 T; 0 U; 0 Other;	
	Query Match	99.9%; Score 3133.4; DB 14; Length 5028;
	Best Local Similarity	99.9%; Pident. No. 0;
	Matches 3134; Conservative	0; Mismatch 1; Indels 0; Gaps 0;
OY	1	ATGGGAGAAAATGAGCAGAGGCCCGACCATTCACGACTCTGGGGGCCCCGGAAAAATTGCG 60
DB	337	ATGGGAGAAAATGAGCAGAGGCCCGACCATTCACGACTCTGGGGGCCCCGGAAAAATTGCG 396
OY	61	GGAATCAAGTACTTTGGAACTGTGAACAGAAGCCGCCCCCAACACCTTCAGAAAACTCGTG 120
DB	397	GGAATCAAGTACTTTGGAACTGTGAACAGAAGCCGCCCCCAACACCTTCAGAAAACTCGTG 456
OY	121	TTCATCAAAAATCCAGATGCTGGTAGAACAACCAGAGGACATTTGAAGTTCCACAAAGAGCT 180
DB	457	TTCATCAAAAATCCAGATGCTGGTAGAACAACCAGAGGACATTTGAAGTTCCACAAAGAGCT 516
OY	181	CCTGGGAAGGTGCTGCTGGTAGAAGCAAGTTTGCACAACCTCAACTCTGTGAAGGTGACTAT 240

Db	517	CTTGAGGAGGCTGCTGGATGCAAGTTTGCAACCACTTCAACCTCGTGAAGTGACAT	576
Qy	241	TTTGGCCCTGAGTTTCTGATCACAACAAAAGATCACGGTGTGGCTGSAATCTCTAAAAACC	300
Db	577	TTTGGCCCTGAGTTTCTGATCACAACAAAAGATCACGGTGTGGCTGSAATCTCTAAAAACC	636
Qy	301	ATTGTGAACAGATTAGAAAGCCAAAGACGTTGTGTTAAGTTTGTGTGAAATTCCTT	360
Db	637	ATTGTGAACAGATTAGAAAGCCAAAGACGTTGTGTTAAGTTTGTGTGAAATTCCTT	696
Qy	361	CCGCTGACCAACACAACTCCCAAGAAATACTCAAGSTACGTTCGCCCTCAGGTG	420
Db	697	CCGCTGACCAACACAACTCCCAAGAAATACTCAAGSTACGTTCGCCCTCAGGTG	756
Qy	421	AAGCAGACTTGGCTCAAGGAGCTTGAAGTGAATGACACACAGCAGCTCTCTTGAT	480
Db	757	AAGCAGACTTGGCTCAAGGAGCTTGAAGTGAATGACACACAGCAGCTCTCTTGAT	816
Qy	481	TCACACATTTGTGCAATCTGAGATTGGGGAATTTGATGAAGCCTTGGACAGAGCACTTA	540
Db	817	TCACACATTTGTGCAATCTGAGATTGGGGAATTTGATGAAGCCTTGGACAGAGCACTTA	876
Qy	541	GCAAAAAATTAATACATCTCTCAGCAAGCGCATAGAGGACAAATCTGTGAAATTTTAC	600
Db	877	GCAAAAAATTAATACATCTCTCAGCAAGCGCATAGAGGACAAATCTGTGAAATTTTAC	936
Qy	601	CATAACCACTTGGCAAAACAACAGACAAATCAAGTTTCCAGCTCTAGAGATTGGCCGT	660
Db	937	CATAACCACTTGGCAAAACAACAGACAAATCAAGTTTCCAGCTCTAGAGATTGGCCGT	996
Qy	661	CGGCTAGAGATGTATGSAATCCGGTTGCACCCGGCCAAAGGACAGGAAGGACAGAAATC	720
Db	997	CGGCTAGAGATGTATGSAATCCGGTTGCACCCGGCCAAAGGACAGGAAGGACAGAAATC	1056
Qy	721	AATCTGCGCTTGGCCAAACGCGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC	780
Db	1057	AATCTGCGCTTGGCCAAACGCGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC	1116
Qy	781	TTCAACTGGGCAAGGTGCGGAGCTGAGCTTCAAGAGGAAGCGCTTCTCATGAAGTC	840
Db	1117	TTCAACTGGGCAAGGTGCGGAGCTGAGCTTCAAGAGGAAGCGCTTCTCATGAAGTC	1176
Qy	841	CGGCGAGATGCAATAGTGCCTACCGAGATACCTTGGAAATCTGATGCGCAGTCGGGAT	900
Db	1177	CGGCGAGATGCAATAGTGCCTACCGAGATACCTTGGAAATCTGATGCGCAGTCGGGAT	1236
Qy	901	TTCTGCAAGTCCCTTGTGAAAATCTGTGTGAATCATAGCCTTCTTGAACCTTTTGA	960
Db	1237	TTCTGCAAGTCCCTTGTGAAAATCTGTGTGAATCATAGCCTTCTTGAACCTTTTGA	1296
Qy	961	GAGGCCAAACCAAAAGCCCAAGCCGCTCTTTAGCGGGGGTATCATTTCCGTTACGT	1020
Db	1297	GAGGCCAAACCAAAAGCCCAAGCCGCTCTTTAGCGGGGGTATCATTTCCGTTACGT	1356
Qy	1021	GGTGCAGACTCAGAAAGAGTTCTCGACTATGTAAAGAGAGGACATTAAGAGGTGCAG	1080
Db	1357	GGTGCAGACTCAGAAAGAGGTCTCGACTATGTAAAGAGAGGACATTAAGAGGTGCAG	1416
Qy	1081	TTTGAAGAGAGCAACAGCAAGATTCTATCCGAGCCTTGCTTCAACACTTACAGAA	1140
Db	1417	TTTGAAGAGAGCAACAGCAAGATTCTATCCGAGCCTTGCTTCAACACTTACAGAA	1476
Qy	1141	CTGAATTCGGAAGTCTGGAGCAAGTCTCAGAGAGACCAACCTTACATTTGGAGAAAGT	1200
Db	1477	CTGAATTCGGAAGTCTGGAGCAAGTCTCAGAGAGACCAACCTTACATTTGGAGAAAGT	1536
Qy	1201	GCCGAATCTCAGAGGGGCGCAAGCTGCGCGGAGAGAAAGAAACGAAGTTTCCGCGGG	1260
Db	1537	GCCGAATCTCAGAGGGGCGCAAGCTGCGCGGAGAGAAAGAAACGAAGTTTCCGCGGG	1596
Qy	1261	GAGCCGGGGTGCACCCGAGCCCTGCGCGGAGAGAAAGCCCGCGGGTTACAGACAGCG	1320

D	b	1597	GAGCCCGGGGTCGACACCCGAGCCCTTCGCGCCGAGAGAAAGCCCGCGGGTAAACAAAGACGGC	1556
O	y	1321	GACCGAGCCCGCTTCGGCGGCCCAACGAGAGAAAGAGAGAGAGTCTGTTAAAGATAGAGACCCAG	1380
D	b	1657	GACGAGCGCGCTTCGGCGGCCCAACGAGAGAAAGAGAGAGAGTCTGTTAAAGATAGAGACCCAG	1716
O	y	1381	CAGAGTAAACTCAGCCCCCGGACCCCAAGACAGAGCTCCCTGACTGGCACTCTCACTT	1440
D	b	1717	CAGAGTAAACTCAGCCCCCGGACCCCAAGACAGAGCTCCCTGACTGGCACTCTCACTT	1776
O	y	1441	TCCGAGCTGTCTGTGAACCTCGCAGGGGGGAGTGCGCCCTCTCAACGAGCACTTGTCTCC	1500
D	b	1777	TCCGAGCTGTCTGTGAACCTCGCAGGGGGGAGTGCGCCCTCTCAACGAGCACTTGTCTCC	1836
O	y	1501	AACCTGAGCCCGGACACCAAGCAGGCTCTCCCTTGAACAAGCCGCTGTGAATGACAG	1560
D	b	1897	AACCTGAGCCCGGACACCAAGCAGGCTCTCCCTTGAACAAGCCGCTGTGAATGACAG	1896
O	y	1561	GCTGCGCCCGGAGCGACATGAGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATAA	1620
D	b	1897	GCTGCGCCCGGAGCGACATGAGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATAA	1956
O	y	1621	GCGATCTTCAATAGCTAAGAGAGTGTACACACCGAGCGGAATATCTGAAAGATCTTGAA	1680
D	b	1957	GCGATCTTCAATAGCTAAGAGAGTGTACACACCGAGCGGAATATCTGAAAGATCTTGAA	2016
O	y	1681	GTTATCACTTCGTGGTTTCAGAGACAGTAGAGCAAAAGAGACGCAATGCCGGAAGCATG	1740
D	b	2017	GTTATCACTTCGTGGTTTCAGAGACAGTAGAGCAAAAGAGACGCAATGCCGGAAGCATG	2076
O	y	1741	AAAAGTCTCATATTCGCCGAATTTTGAACCTTTGCAAAATTTCAACTAATTTTCTCAAG	1800
D	b	2077	AAAAGTCTCATATTCGCCGAATTTTGAACCTTTGCAAAATTTCAACTAATTTTCTCAAG	2136
O	y	1801	GAAATTGAGGACGACCTTGCCCTGTGGGAAAGGCGCTCAAAATGCCAAATCAGAGATTAC	1860
D	b	2137	GAAATTGAGGACGACCTTGCCCTGTGGGAAAGGCGCTCAAAATGCCAAATCAGAGATTAC	2196
O	y	1861	CAAAAGATTCGGCCGATGTCAATGCTGAGAAACAATTCAGGGCAATGAAGCACTTCGGCGCTCAC	1920
D	b	2197	CAAAAGATTCGGCCGATGTCAATGCTGAGAAACAATTCAGGGCAATGAAGCACTTCGGCGCTCAC	2256
O	y	1921	CTGTGGAAGCACAGCGAGGCTTTGAGGCGCTTGAGAGATGGAATCAAGACCTCCCGCGG	1980
D	b	2257	CTGTGGAAGCACAGCGAGGCTTTGAGGCGCTTGAGAGATGGAATCAAGACCTCCCGCGG	2316
O	y	1981	CTGAGAGAACTTTCGCAAGACCTTTGAGCTCAGAGGTGTGTTAATCTAACCGCTCAACACC	2040
D	b	2317	CTGAGAGAACTTTCGCAAGACCTTTGAGCTCAGAGGTGTGTTAATCTAACCGCTCAACACC	2376
O	y	2041	TTCTCTCTGCGGCACTGTCAACCGGCTCATCTCACTAACAGAGGTCCTGAGACGGCTGTGC	2100
D	b	2377	TTCTCTCTGCGGCACTGTCAACCGGCTCATCTCACTAACAGAGGTCCTGAGACGGCTGTGC	2436
O	y	2101	AAACACCAACCCCGCGAGCCACGCGCACTTCAGGAGCTGCGAGCCGCTTTGGCAGAGATC	2160
D	b	2437	AAACACCAACCCCGCGAGCCACGCGCACTTCAGGAGCTGCGAGCCGCTTTGGCAGAGATC	2496
O	y	2161	ACGAGATGTGTGGACAGCTTCACACGATACATATCAAGATGAGAAATTTCCAGAAAGCTG	2220
D	b	2497	ACGAGATGTGTGGACAGCTTCACACGATACATATCAAGATGAGAAATTTCCAGAAAGCTG	2556
O	y	2221	CACGAACTCAAGAAAGATTTGATTGGCATTTGATTTGTGGTCCGGGAAAGGAGATTTC	2280
D	b	2557	CACGAACTCAAGAAAGATTTGATTGGCATTTGATTTGTGGTCCGGGAAAGGAGATTTC	2616
O	y	2281	ATTCGCTGTGGGACAGCTTCAGCAAGCTTCGCGGGAAGGGGCTCCAGCAGCGCATGTTCTC	2340
D	b	2617	ATTCGCTGTGGGACAGCTTCAGCAAGCTTCGCGGGAAGGGGCTCCAGCAGCGCATGTTCTC	2676
O	y	2341	CTGTTCAACGACGTCTGTCTATACAGAGCCGGGGGCTGAACGGCTTCGAATCAGTTTAA	2400
D	b	2677	CTGTTCAACGACGTCTGTCTATACAGAGCCGGGGGCTGAACGGCTTCGAATCAGTTTAA	2736

```
QY 2401 GTCCACGGGGGAGCTCCGCTCTATGGAATGACGATTGAGAGAGGAAAGCAAGTGGGGG 2460
DB 2737 GTCCACGGGGGAGCTCCGCTCTATGGAATGAGATGAGAGAGCAAGCAAGTGGGGG 2796
QY 2461 GTGCCCCACCTGCTGACCTCTCCGGGGCCAGCGGAGTCCATCATCTGAGCGCCGCACTTCT 2520
DB 2797 GTGCCCCACCTGCTGACCTCTCCGGGGCCAGCGGAGTCCATCATCTGAGCGCCGCACTTCT 2856
QY 2521 CGGTCCGAGATGAGAAAGTGGGTGAGAGCATTCAGATGACCATTTGACTTGGCCGAGAG 2580
DB 2857 CGGTCCGAGATGAGAAAGTGGGTGAGAGCATTCAGATGACCATTTGACTTGGCCGAGAG 2916
QY 2581 AGCAGACGCCGCCGCTGAGTTCTGAGCCAGACGCCGCCGTCACAAAGTCCCTGAT 2640
DB 2917 AGCAGACGCCGCCGCTGAGTTCTGAGCCAGACGCCGCCGTCACAAAGTCCCTGAT 2976
QY 2641 GAAAGCCACCGCGGCTGACAGAGATCAGAGATGACCTGAGCGCTCGCGCACATCGCTG 2700
DB 2977 GAAAGCCACCGCGGCTGACAGAGATCAGAGATGACCTGAGCGCTCGCGCACATCGCTG 3036
QY 2701 GAGCGCCAGGCGCCGCGCACCGCGGCACACAAATGATGACCTGCTGCGCACCGCACACC 2760
DB 3037 GAGCGCCAGGCGCCGCGCACCGCGGCACACAAATGATGACCTGCTGCGCACCGCACACC 3096
QY 2761 AGGCTCTCAATGATGACCTTCAAGATGAGAGATGAGATGATGAGAAACCTGCTG 2820
DB 3097 AGGCTCTCAATGATGACCTTCAAGATGAGAGATGAGATGATGAGAAACCTGCTG 3156
QY 2821 AGGAAATTTCAAAAACAGCAACGCGGTGCGAGAGCTGTGGGTGTTCACAACTTCTGC 2880
DB 3157 AGGAAATTTCAAAAACAGCAACGCGGTGCGAGAGCTGTGGGTGTTCACAACTTCTGC 3216
QY 2881 CTGTTCTTTCACAAATACACACAGAGCAATCATCTCCCTGCGACCTGCTGCTGCGC 2940
DB 3217 CTGTTCTTTCACAAATACACACAGAGCAATCATCTCCCTGCGACCTGCTGCTGCGC 3276
QY 2941 TACTCGCTACACATCCCTCTGAGTCCGAGAAACATTCAGAAAGCTACGTTTCAGACTG 3000
DB 3277 TACTCGCTACACATCCCTCTGAGTCCGAGAAACATTCAGAAAGCTACGTTTCAGACTG 3336
QY 3001 CACTTCAGATCCGACGCTCTACTACTTCAGGGCGGAAAGCAGATACAGTTTCAAAAGTGG 3060
DB 3337 CACTTCAGATCCGACGCTCTACTACTTCAGGGCGGAAAGCAGATACAGTTTCAAAAGTGG 3396
QY 3061 ATGGAAGTATCCGACAGTGCACCAAGCTGTGCTCGGACCCCACTGTTGAAGCCACAA 3120
DB 3397 ATGGAAGTATCCGACAGTGCACCAAGCTGTGCTCGGACCCCACTGTTGAAGCCACAA 3456
QY 3121 GAGTCTCTTGTGTAT 3135
DB 3457 GAGTCTCTTGTGTAT 3471

RESULT 4
AAA08582
ID AAA08582 standard; DNA; 4687 BP.
XX
XX AAA08582;
XX
XX 19-JUL-2000 (first entry)
XX
DE Human cytoskeleton associated protein 2 (CYSKP-2) coding sequence.
XX
XX Cytoskeleton associated protein; CYSKP-2; cancer; proliferative;
XX autoimmunity; inflammatory; vesicle trafficking; neurological;
XX cardiovascular; cell motility; reproductive; muscle disorder; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 90..3227
XX FT /*tag= a
```

```
FT XX /product= "CYSKP-2"
XX PN W0200017355-A2.
XX PD 30-MAR-2000.
XX PE 17-SEP-1999; 99MO-US021565.
XX PF 18-SEP-1998; 98US-0172226P.
XX PR 27-APR-1999; 99US-0131321P.
XX PA (INCY-) INCYTE PHARM INC.
XX
XX Lel P, Tang YT, Yue H, Hillman JL, Bandman O, Corley NC;
XX Guegler KJ, Patterson C, Azimzai Y, Baughn MR;
XX WPI; 2000-283582/24.
XX DR P-PSDB; AAY91947.
XX
XX Human cytoskeleton associated proteins, used to treat cell proliferative,
XX autoimmunity/inflammatory, vesicle trafficking, neurological, cell
XX motility, reproductive and muscle disorders.
XX
XX Claim 9; Page 101-102; 113pp; English.
XX
XX AAA08581-96 encode human cytoskeleton associated proteins 1 to 16 (CYSKP-
XX 1 to CYSKP-16) respectively. The sequences can be used to treat and
XX diagnose cancer and cell proliferative, autoimmunity/inflammatory, vesicle
XX trafficking, neurological, cardiovascular, cell motility, reproductive
XX and muscle disorders. Pharmaceutical compositions containing CYSKP-2 to
XX CYSKP-16 can be used to treat or prevent disorders associated with
XX decreased expression or activity of CYSKP (claimed), for example,
XX atherosclerosis, cirrhosis, hepatitis, myelofibrosis, psoriasis, cancer,
XX autoimmune/antiinflammatory disorders such as allergies, anemia, asthma,
XX acquired immunodeficiency syndrome (AIDS), Crohn's disease, diabetes
XX mellitus, Goodpasture's syndrome, osteoarthritis, scleroderma and trauma.
XX CYSKP antagonists can be used to treat or prevent a disorder associated
XX with increased expression or activity of CYSKP (claimed)
XX
XX Sequence 4687 BP; 1179 A; 1261 C; 1196 G; 1051 T; 0 U; 0 Other;
XX
XX Query Match 99.9%; Score 3131.8; DB 3; Length 4687;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 3133; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGAGAAATATAGACAGAGCGGACCCGACATCAGACTGGGGGCCCGGAAATTCG 60
DB 90 ATGGAGAAATATAGACAGAGCGGACCCGACATCAGACTGGGGGCCCGGAAATTCG 149
QY 61 GGGATCAGTACCTTGGAAAGTGGACAGAAAGCCGCCCAACACTTCAGAGAAATCTG 120
DB 150 GGGATCAGTACCTTGGAAAGTGGACAGAAAGCCGCCCAACACTTCAGAGAAATCTG 209
QY 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCAATTTGAAGTTCACAAAGAGCT 180
DB 210 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCAATTTGAAGTTCACAAAGAGCT 269
QY 210 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCAATTTGAAGTTCACAAAGAGCT 269
DB 270 CTTGGAGAGTGTCTGATGATGACAGTTTCAACACCTCAACTCTGTGAGAGTGA 240
QY 181 CTTGGAGAGTGTCTGATGATGACAGTTTCAACACCTCAACTCTGTGAGAGTGA 240
DB 270 CTTGGAGAGTGTCTGATGATGACAGTTTCAACACCTCAACTCTGTGAGAGTGA 240
QY 241 TTTGGCTTGAAGTTTCTGATGACAAAGAAATGACGCTGTGATCTCTCAAAAGCC 300
DB 330 TTTGGCTTGAAGTTTCTGATGACAAAGAAATGACGCTGTGATCTCTCAAAAGCC 389
QY 301 ATTTGAAACAGATTGAAGGCCAAAGCAAGTTGTTGAAGTTTGAAGTTTGAAGTTTCT 360
DB 390 ATTTGAAACAGATTGAAGGCCAAAGCAAGTTGTTGAAGTTTGAAGTTTGAAGTTTCT 449
QY 361 CCGCTGACCAACACAACTCCAAAGAACTTCAAGAGTACCTGTTGCGCTGACAGTG 420
DB 450 CCGCTGACCAACACAACTCCAAAGAACTTCAAGAGTACCTGTTGCGCTGACAGTG 509
```

421 AAGCAGAGCTTGCTCAAGGAGTTGACGTGTAATGACACAGCGAGCTCTTGATT 480
| | | | |
Db 510 AAGCAGAGCTTGCTCAAGGAGTTGACGTGTAATGACACAGCGAGCTCTTGATT 569
| | | | |
Qy 481 TCACACATTTGTCAATCTGAGATTTGGGAAATTTGATGAAGCTTGGACAGAGACCTTA 540
| | | | |
Db 570 TCACACATTTGTCAATCTGAGATTTGGGAAATTTGATGAAGCTTGGACAGAGACCTTA 629
| | | | |
Qy 541 GGAAGAAATTAATTAATCTGACGAGAGAGCGACTGAGAGACAAATTCGTGAATTTGAC 600
| | | | |
Db 630 GGAAGAAATTAATTAATCTGACGAGAGAGCGACTGAGAGACAAATTCGTGAATTTGAC 689
| | | | |
Qy 601 CATTAACCATTTGAGACAAACACAGCAGATCAGATTTCCAGCTCTTAGAGATTCGCCGT 660
| | | | |
Db 690 CATTAACCATTTGAGACAAACACAGCAGATCAGATTTCCAGCTCTTAGAGATTCGCCGT 749
| | | | |
Qy 661 CGGCTAGAGATGTATGAAATCCGGTTGCAACCGGCCAAGAGACAGGAGGACAGAAATC 720
| | | | |
Db 750 CGGCTAGAGATGTATGAAATCCGGTTGCAACCGGCCAAGAGACAGGAGGACAGAAATC 809
| | | | |
Qy 721 AATCTGGCCGTTGGCCCAACCGGGAATTTCTAGTTTCAAGGTTTCACTAAGATCAATGCG 780
| | | | |
Db 810 AATCTGGCCGTTGGCCCAACCGGGAATTTCTAGTTTCAAGGTTTCACTAAGATCAATGCG 869
| | | | |
Qy 781 TTCAACTGGGCCAAGGTGCGAAAGCTGAGCTTCAAGAGAGAGCGCTTCTCAATCAAGCTC 840
| | | | |
Db 870 TTCAACTGGGCCAAGGTGCGAAAGCTGAGCTTCAAGAGAGAGCGCTTCTCAATCAAGCTC 929
| | | | |
Qy 841 CGGCCAGATGCCAATATGTGCTTACAGGATACCTTGAATTCCTGATGCGCAGTCCGAGAT 900
| | | | |
Db 930 CGGCCAGATGCCAATATGTGCTTACAGGATACCTTGAATTCCTGATGCGCAGTCCGAGAT 989
| | | | |
Qy 901 TTCTGAAATCTCTTGGGAAATCTGTGTGAACATCAGCTCTTCTTGAATTTTGA 960
| | | | |
Db 990 TTCTGAAATCTCTTGGGAAATCTGTGTGAACATCAGCTCTTCTTGAATTTTGA 1049
| | | | |
Qy 961 GAGCCCAACCAAGGCCAGCCGTCCTTTAGCCGGGGGTCAATCTTGGTTCAGT 1020
| | | | |
Db 1050 GAGCCCAACCAAGGCCAGCCGTCCTTTAGCCGGGGGTCAATCTTGGTTCAGT 1109
| | | | |
Qy 1021 GGTGCGACTCAGAGAGAGGTTCTGACTATGTTAAAGAGAGACATTAAGAGGTGCG 1080
| | | | |
Db 1110 GGTGCGACTCAGAGAGAGGTTCTGACTATGTTAAAGAGAGACATTAAGAGGTGCG 1169
| | | | |
Qy 1081 TTGTAAGAGAGAGAGAGAGGTTCAATTTATCCGAGCTTGTCAAGCTTACAGAA 1140
| | | | |
Db 1170 TTGTAAGAGAGAGAGAGGTTCAATTTATCCGAGCTTGTCTTCAAGCTTACAGAA 1229
| | | | |
Qy 1141 CTGAATTCGAAAGTGTGAGAGAGTCTCAGCAGAGACAGCCTTACATTTGGAGAGGT 1200
| | | | |
Db 1230 CTGAATTCGAAAGTGTGAGAGAGTCTCAGCAGAGACAGCCTTACATTTGGAGAGGT 1289
| | | | |
Qy 1201 GCCGAATCTCAGAGGGGCCAGAGCTCCGCGAGAGAAAGAACCGAAGTTTCCCGCGG 1260
| | | | |
Db 1290 GCCGAATCTCAGAGGGGCCAGAGCTCCGCGAGAGAAAGAACCGAAGTTTCCCGCGG 1349
| | | | |
Qy 1261 GAGCCGGGGTGGACCCGAGGCCCTGCGCGAGAGAGAACCCCGCGGGTAAACAAGAGGGG 1320
| | | | |
Db 1350 GAGCCGGGGTGGACCCGAGGCCCTGCGCGAGAGAGAACCCCGCGGGTAAACAAGAGGGG 1409
| | | | |
Qy 1321 GACGAGAGCGGCTCGAGCGGCCACGAGAGAGAGAGAGAGGTCTGTTAAGATAGAACCCAG 1380
| | | | |
Db 1410 GACGAGAGCGGCTCGAGCGGCCACGAGAGAGAGAGAGGTCTGTTAAGATAGAACCCAG 1469
| | | | |
Qy 1381 CAGAGTAAACTCTAGCCCCCGCAGCCAGACAGGCTCCCTGACTGGCAGTCTTCACTTT 1440
| | | | |
Db 1470 CAGAGTAAACTCTAGCCCCCGCAGCCAGACAGGCTCCCTGACTGGCAGTCTTCACTTT 1529
| | | | |
Qy 1441 TCCGAGCTGTCTGTGAATCTGCAAGGGGGAGTGGCCCTGCGCAAGCTTGAATTTGCC 1500
| | | | |
Db 1530 TCCGAGCTGTCTGTGAATCTGCAAGGGGGAGTGGCCCTGCGCAAGCTTGAATTTGCC 1589
| | | | |
Qy 1501 AACCTGAGCCCGCAGACCAAGCAGGCTCTCCCTTGATCAGCCCGCTGTGATGACAG 1560
| | | | |

1590 AACCTGAGCCCGCAGACCAAGAGGCTCTCCCTTGATCAGCCCGCTGTGATGACAG 1649
| | | | |
Qy 1561 GCTTCCCCCGGACGGAAGATGAGAGTGAAGGGCCGGAAGAAAGATTTCCCACTGAATAA 1620
| | | | |
Db 1650 GCTTCCCCCGGACGGAAGATGAGAGTGAAGGGCCGGAAGAAAGATTTCCCACTGAATAA 1709
| | | | |
Qy 1621 GCGTACTCATATGACTAAGAGATGTCTACACCGAGCGCAATATCTGAAGATCTCGAA 1680
| | | | |
Db 1710 GCGTACTCATATGACTAAGAGATGTCTACACCGAGCGCAATATCTGAAGATCTCGAA 1769
| | | | |
Qy 1681 GTTATCACTGCTGCTTTCAAGACACAGTGAAGCAAGAGAGCGCATCCGGAACACTG 1740
| | | | |
Db 1770 GTTATCACTGCTGCTTTCAAGACACAGTGAAGCAAGAGAGCGCATCCGGAACACTG 1829
| | | | |
Qy 1741 AAAAGTCTCATATTTCCGGAATTTTGAACCTTTGCAAAATTTCAATCTAATTTTCTCAAG 1800
| | | | |
Db 1830 AAAAGTCTCATATTTCCGGAATTTTGAACCTTTGCAAAATTTCAATCTAATTTTCTCAAG 1889
| | | | |
Qy 1801 GAAATTGAGCAACGACTTGGCCCTGTGGAGAGGCGCTCAAAATGCCCAATCAGATTAAC 1860
| | | | |
Db 1890 GAAATTGAGCAACGACTTGGCCCTGTGGAGAGGCGCTCAAAATGCCCAATCAGATTAAC 1949
| | | | |
Qy 1861 CAAGAAATCGGAGATGTATGCTGAAGAAATTCAGGGCATGAAGCACCTGGCGGCTCAC 1920
| | | | |
Db 1950 CAAGAAATCGGAGATGTATGCTGAAGAAATTCAGGGCATGAAGCACCTGGCGGCTCAC 2009
| | | | |
Qy 1921 CTGTGGAAGCACAGGAGGCGCTTGAAGGCGCTTGAAGAAATGGAATGAAGATCCCGGGCG 1980
| | | | |
Db 2010 CTGTGGAAGCACAGGAGGCGCTTGAAGGCGCTTGAAGAAATGGAATGAAGATCCCGGGCG 2069
| | | | |
Qy 1981 CTGAGAACTTCTGAGAGACTTTGAGCTGAGAGAGAGTGTGTAACCTGCTCAACACC 2040
| | | | |
Db 2070 CTGAGAACTTCTGAGAGACTTTGAGCTGAGAGAGAGTGTGTAACCTGCTCAACACC 2129
| | | | |
Qy 2041 TTCTCTGCGGCACTGACCGGCTCATGCACTAACAGAGCTCTTGAAGCGGCTGTGTC 2100
| | | | |
Db 2130 TTCTCTGCGGCACTGACCGGCTCATGCACTAACAGAGCTCTTGAAGCGGCTGTGTC 2189
| | | | |
Qy 2101 AAACACACCCCGCGAGCAGCGCCGACTTTCAGGGGCTGCGAGCGCTTGGCAGAGATC 2160
| | | | |
Db 2190 AAACACACCCCGCGAGCAGCGCCGACTTTCAGGGGCTGCGAGCGCTTGGCAGAGATC 2249
| | | | |
Qy 2161 ACGAGATGTGTGACACAGCTCCACGCTACAGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2220
| | | | |
Db 2250 ACGAGATGTGTGACACAGCTCCACGCTACAGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2309
| | | | |
Qy 2221 CACGAATCTAAGAAAGATTTGATTGGCAATTGCAATCTTTGTGTTCCGGAAAGGAGTTTC 2280
| | | | |
Db 2310 CACGAATCTAAGAAAGATTTGATTGGCAATTGCAATCTTTGTGTTCCGGAAAGGAGTTTC 2369
| | | | |
Qy 2281 ATCCGCTGTGGGAGGCTGACGAGAGCTCTCGGGGAAAGGGGCTCCAGCAGGCAATGTTCTTC 2340
| | | | |
Db 2370 ATCCGCTGTGGGAGGCTGACGAGAGCTCTCGGGGAAAGGGGCTCCAGCAGGCAATGTTCTTC 2429
| | | | |
Qy 2341 CTGTTCAACGAGAGTCTGCTCTATACAGAGCGGGGCTGACGCGCTCCAAATCAGTTTAAA 2400
| | | | |
Db 2430 CTGTTCAACGAGAGTCTGCTCTATACAGAGCGGGGCTGACGCGCTCCAAATCAGTTTAAA 2489
| | | | |
Qy 2401 GTTCACGGGCAAGCTCCCGCTCTATGAGATGACGATTTGAGAGAGCGAAAGACAGATGGGGG 2460
| | | | |
Db 2490 GTTCACGGGCAAGCTCCCGCTCTATGAGATGACGATTTGAGAGAGCGAAAGACAGATGGGGG 2549
| | | | |
Qy 2461 GTTCCCACTGCTGACCTCCCGGGGCAAGGGGCAAGTCCATCATGTGTGGCGCCAGTTTCT 2520
| | | | |
Db 2550 GTTCCCACTGCTGACCTCCCGGGGCAAGGGGCAAGTCCATCATGTGTGGCGCCAGTTTCT 2609
| | | | |
Qy 2521 CGGTCCAGATGAGAGAGTGGGTTGAGACATTCAGATGAGCAATGACCTTGGCGGAGAG 2580
| | | | |
Db 2610 CGGTCCAGATGAGAGAGTGGGTTGAGACATTCAGATGAGCAATGACCTTGGCGGAGAG 2669
| | | | |
Qy 2581 AGCAGAGCCCGCCCTGAGTTCTTGGCCAGCAGCCCTCTGACACAGATGCCCTGAT 2640
| | | | |

Db	2670	AGCAGCAGACCCCGCCCGCTGAGTTCTCTGGCCAGACAGCCCCCTTGACACAAAGTCCCCAT	2723
Qy	2641	GAAGCCACCGCGGCTGACACAGAGTCAGAGGATGACTTGAGCGCCTCGCGCAATCGCTG	2700
Db	2730	GAAGCCACCGCGGCTGACACAGAGTCAGAGGATGACTTGAGCGCCTCGCGCAATCGCTG	2789
Qy	2701	GAGGCGCCAGGCGCCCGGCAACCGCGGCAACCAATGTGTGCATGTGTCTGGCAACGCAACC	2766
Db	2790	GAGGCGCCAGGCGCCCGGCAACCGCGGCAACCAATGTGTGCATGTGTCTGGCAACGCAACC	2849
Qy	2761	AGCGCTCCATGTGTGACTTCAGCATGCGACGTGAGGATCAAGTTGTCTGGAAACCTGCTG	2820
Db	2850	AGCGCTCCATGTGTGACTTCAGCATGCGACGTGAGGATCAAGTTGTCTGGAAACCTGCTG	2909
Qy	2821	AGGAAATTCAAAACAGCAACCGGCTGGCAGAAAGCTGTGGTGTGTTCACAAATTTCTGC	2880
Db	2910	AGGAAATTCAAAACAGCAACCGGCTGGCAGAAAGCTGTGGTGTGTTCACAAATTTCTGC	2965
Qy	2881	CTGTTCTTTCTTCAAAATCACACCAAGCAATATCCCTTGGCCAGCTGCTCTGCTCGGC	2940
Db	2970	CTGTTCTTTCTTCAAAATCACACCAAGCAATATCCCTTGGCCAGCTGCTCTGCTCGGC	3029
Qy	2941	TACTCGCTCACCAATCCCTCTGAGTCCGAGAAACATCCGAAAGACTACGTGTTCAAGCTG	3000
Db	3030	TACTCGCTCACCAATCCCTCTGAGTCCGAGAAACATCCGAAAGACTACGTGTTCAAGCTG	3089
Qy	3001	CAC TTC AAGTCCCAAGTCTACTCTTCAGGGCGGAAACAGATGACACGTTGAAAGGTGG	3066
Db	3090	CAC TTC AAGTCCCAAGTCTACTCTTCAGGGCGGAAACAGATGACACGTTGAAAGGTGG	3149
Qy	3061	ATGGAAGTGAATCCGCAAGTCCACCAACGCTCTGCGACACCCCAAGTGTGAGCCACAA	3120
Db	3150	ATGGAAGTGAATCCGCAAGTCCACCAACGCTCTGCGACACCCCAAGTGTGAGCCACAA	3209
Qy	3121	GAGTCTCTGTGTGTAT	3135
Db	3210	GAGTCTCTGTGTAT	3224

RESULT	
ID	AAS64760 standard; cDNA, 3187 BP.
AC	AAS64760;
DT	13-FEB-2002 (first entry)
DE	DNA encoding novel human diagnostic protein #564.
KW	Human; chromosome mapping; gene mapping; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss
OS	Homo sapiens.
PN	MO200175067-A2.
PD	11-OCT-2001.
Pf	30-MAR-2001; 2001MO-US008631.
PR	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649267.
PA	(HYSE-) HYSEQ INC.
PI	Dymanac RT, Liu C, Tang YT;
PP	
DR	WPI; 2001-639362/73. P-PsDB; ABG00573.
PT	New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess

Query Match	Best Local Similarity	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
1	99.5%	3119.2	5	3187	1	0	3	1	1
49	99.9%	3119.2	5	3187	1	0	3	1	1
61	99.9%	3119.2	5	3187	1	0	3	1	1
109	99.9%	3119.2	5	3187	1	0	3	1	1
121	99.9%	3119.2	5	3187	1	0	3	1	1
169	99.9%	3119.2	5	3187	1	0	3	1	1
181	99.9%	3119.2	5	3187	1	0	3	1	1
229	99.9%	3119.2	5	3187	1	0	3	1	1
241	99.9%	3119.2	5	3187	1	0	3	1	1
289	99.9%	3119.2	5	3187	1	0	3	1	1
301	99.9%	3119.2	5	3187	1	0	3	1	1
349	99.9%	3119.2	5	3187	1	0	3	1	1
361	99.9%	3119.2	5	3187	1	0	3	1	1
409	99.9%	3119.2	5	3187	1	0	3	1	1
421	99.9%	3119.2	5	3187	1	0	3	1	1
469	99.9%	3119.2	5	3187	1	0	3	1	1
481	99.9%	3119.2	5	3187	1	0	3	1	1
529	99.9%	3119.2	5	3187	1	0	3	1	1
541	99.9%	3119.2	5	3187	1	0	3	1	1
589	99.9%	3119.2	5	3187	1	0	3	1	1
601	99.9%	3119.2	5	3187	1	0	3	1	1

|||||
Db 649 CATPACCAATTGGAAACACACAGAGATCATGATTTCCAGTCTCTAAGATTGCCGT 708
Qy 661 CGGCTAGAGATGTTATGGAAATCCGGTTGCACCCGGCCAAAGACAGGGAAGCCAGGAAGATC 720
Db 709 CGGCTAGAGATGTTATGGAAATCCGGTTGCACCCGGCCAAAGACAGGGAAGCCAGGAAGATC 768
Qy 721 AATCTGGCCGTTGGCCAAACCGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGTC 780
Db 769 AATCTGGCCGTTGGCCAAACCGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGTC 828
Qy 781 TTCAACTGGGCCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGCTTTCTCACTCAAGTTC 840
Db 829 TTCAACTGGGCCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGCTTTCTCACTCAAGTTC 888
Qy 841 CGGCCAGATGCCAATATGTCGTAACAGAGATCCTTGGAAATTCCTGATGCGCAGTCCGGAT 900
Db 889 CGGCCAGATGCCAATATGTCGTAACAGAGATCCTTGGAAATTCCTGATGCGCAGTCCGGAT 948
Qy 901 TTCTGCAAGTCTTCTGGAATAATCTGTGTGAACATCATGCTTCTTTAGACTTTTGA 960
Db 949 TTCTGCAAGTCTTCTGGAATAATCTGTGTGAACATCATGCTTCTTTAGACTTTTGA 1008
Qy 961 GAGCCCAACCAAAAGCCCAAGCCCGTCTCTTTAGCCGGGGGTCAATTTTGGTTCACT 1020
Db 1009 GAGCCCAACCAAAAGCCCAAGCCCGTCTCTTTAGCCGGGGGTCAATTTTGGTTCACT 1068
Qy 1021 GGTCCGACTCAGAAAGAGGTTCTCCACTATGTTAAAGAAAGAGACATAAGAAAGGTGCAG 1080
Db 1069 GGTCCGACTCAGAAAGAGGTTCTCCACTATGTTAAAGAAAGAGACATAAGAAAGGTGCAG 1128
Qy 1081 TTTGAAAGGAGACAGCAAGATTCAATTCTATCCGAGCCTTGCTTCAAGCCTACAGAA 1140
Db 1129 TTTGAAAGGAGACAGCAAGATTCAATTCTATCCGAGCCTTGCTTCAAGCCTACAGAA 1188
Qy 1141 CTGAATTCGAAAGTGTCTGAGAGCTCTCAGCAGAGACCAAGCCTTACATTTGGAGA -AGG 1199
Db 1189 CTGAATTCGAAAGTGTCTGAGAGCTCTCAGCAGAGACCAAGCCTTACATTTGGAGAAGG 1248
Qy 1200 TGGCCGATCTCCAGGGGGGCCAAGCTGCGCGGAGAAAGAAACGGAAGTTTCCGCCGG 1259
Db 1249 TGGCCGATCTCCAGGGGGGCCAAGCTGCGCGGAGAAAGAAACGGAAGTTTCCGCCGG 1308
Qy 1260 GGAAGCCGGGGTGCACCCGAGCCCTGCGCGAGGAAGACCCCGGGGTAAACGAAGAGGC 1319
Db 1309 GGAAGCCGGGGTGCACCCGAGCCCTGCGCGAGGAAGACCCCGGGGTAAACGAAGAGGC 1368
Qy 1320 GGAAGAGCCGCTCGCGCCCAAGAGGAAGAGAGAGTCTGTTAAGATATGAGACCCA 1379
Db 1369 GGAAGAGCCGCTCGCGCCCAAGAGGAAGAGAGAGTCTGTTAAGATATGAGACCCA 1428
Qy 1380 GCAAGATTAACCTCAGCCCCCGCAGCCAGACAGACAGGCTCTGATCGGACATCTCACT 1439
Db 1429 GCAAGATTAACCTCAGCCCCCGCAGCCAGACAGACAGGCTCTGATCGGACATCTCACT 1488
Qy 1440 TTCGAGCTGTCTGTAACCTCGCAGGGGGAGTGGCCCTGCAACGTAACCTTGTCTCC 1499
Db 1489 TTCGAGCTGTCTGTAACCTCGCAGGGGGAGTGGCCCTGCAACGTAACCTTGTCTCC 1548
Qy 1500 CAACCTGAGCCCCGACACCAAGAGGCTCTCCCTTGTATCAAGCCCGCTGTAATGACA 1559
Db 1549 CAACCTGAGCCCCGACACCAAGAGGCTCTCCCTTGTATCAAGCCCGCTGTAATGACA 1608
Qy 1560 GGCCTTCGCCCCCGGACGGAATGAGATGAGGGCCGGAAGAAAGATTTCCCACTGATTA 1619
Db 1609 GGCCTTCGCCCCCGGACGGAATGAGATGAGGGCCGGAAGAAAGATTTCCCACTGATTA 1668
Qy 1620 AGCGTACTTCATAGCTAAGAAAGTGTCTACCAACGAGCGAACAATATCTGAAGATCTCGA 1679
Db 1669 AGCGTACTTCATAGCTAAGAAAGTGTCTACCAACGAGCGAACAATATCTGAAGATCTCGA 1728
Qy 1680 AGTTATCACTTGTGTGTTTCAAGACAAGTGAACAAAGAGAGCCATGCCGGAAGCACT 1739
|||||

Db 1729 AGTTATCACTTGTGTGTTTCAAGACACAGTGAACAAAGAGACGCCATGCCGAAGCACT 1788
Qy 1740 GAAAAGTCTCATATCCCGAATTTTGAACCTTTGACAAATTTCTACTAATTTTCTCA 1799
Db 1789 GAAAAGTCTCATATCCCGAATTTTGAACCTTTGACAAATTTCTACTAATTTTCTCA 1848
Qy 1800 GAAAATTTAGCAACGACTTGGCTGTGGAAAGCCGCTCAATATGCCCAATCAGAGATTA 1859
Db 1849 GAAAATTTAGCAACGACTTGGCTGTGGAAAGCCGCTCAATATGCCCAATCAGAGATTA 1908
Qy 1860 CCAAGAAATCCGCGATGTCTGTAAGAAACATTCAGGCAATGAAGCACTGGCCGCTCA 1919
Db 1909 CCAAGAAATCCGCGATGTCTGTAAGAAACATTCAGGCAATGAAGCACTGGCCGCTCA 1968
Qy 1920 CCTGTGGAAGCAACGCGAGGCTTGGAGGCTTGGAGATGGAATCAAGAGCTCCGGCG 1979
Db 1969 CCTGTGGAAGCAACGCGAGGCTTGGAGGCTTGGAGATGGAATCAAGAGCTCCGGCG 2028
Qy 1980 GCTGGAAGAACTTCTGACAGACTTTGAGCTGCAGAAAGGTGTGTACTTACCTACCGCTCAAC 2039
Db 2029 GCTGGAAGAACTTCTGACAGACTTTGAGCTGCAGAAAGGTGTGTACTTACCTACCGCTCAAC 2088
Qy 2040 CTTCCTCTGCGGCCACTGTCACCGGCTCATGCACTAACAGAGTCTGAGCGGCTGTG 2099
Db 2089 CTTCCTCTGCGGCCACTGTCACCGGCTCATGCACTAACAGAGTCTGAGCGGCTGTG 2148
Qy 2100 CAAACACCAACCGCGAGCAACGCGACTTCAGGGGCTGCCAGCGGCTTTGGCAGAGAT 2159
Db 2149 CAAACACCAACCGCGAGCAACGCGACTTCAGGGGCTGCCAGCGGCTTTGGCAGAGAT 2208
Qy 2160 CACGAGATGTGTGACAGACTTCAAGGTACATGATCAAGATGGAATAATTTCCAGAGCT 2219
Db 2209 CACGAGATGTGTGACAGACTTCAAGGTACATGATCAAGATGGAATAATTTCCAGAGCT 2268
Qy 2220 GCACGAATCTGAAGAAATTTGATTTGGCAATTGACATCTTGTGTTCCGGGAAGGAGTT 2279
Db 2269 GCACGAATCTGAAGAAATTTGATTTGGCAATTGACATCTTGTGTTCCGGGAAGGAGTT 2328
Qy 2280 CATCCGTCTGGGCAACCTCAGGAAGCTCTGGGGGAAGGGGCTCCAGACGCGCATGTTCTT 2339
Db 2329 CATCCGTCTGGGCAACCTCAGGAAGCTCTGGGGGAAGGGGCTCCAGACGCGCATGTTCTT 2388
Qy 2340 CCTGTTCAACGACGTCCTGCTATACACGAGCCGGGGGCTGACGGCCTCCAACTCACTTAA 2399
Db 2389 CCTGTTCAACGACGTCCTGCTATACACGAGCCGGGGGCTGACGGCCTCCAACTCACTTAA 2448
Qy 2400 AGTCCACGGGAGCTCCGCTCTATGCACTGAACGATTAAGAGAGACGAAGAGTGGGG 2459
Db 2449 AGTCCACGGGAGCTCCGCTCTATGCACTGAACGATTAAGAGAGACGAAGAGTGGGG 2508
Qy 2460 GGTGCCCCAATGCTGACCTCCGGGGGCAAGGGGAGTCATCATGCTGGCCGCAATTTC 2519
Db 2509 GGTGCCCCAATGCTGACCTCCGGGGGCAAGGGGAGTCATCATGCTGGCCGCAATTTC 2568
Qy 2520 TCGGTCCGAGATGGAAGATGGGTTGAGAGACATCCAGATGGCCATTGACCTGGCGAGAA 2579
Db 2569 TCGGTCCGAGATGGAAGATGGGTTGAGAGACATCCAGATGGCCATTGACCTGGCGAGAA 2628
Qy 2580 GAGCAGCAGCCCCCGCTGAGTTCCTGGCAGACAGCCCCCTGCAACAAGTCCCCCTGA 2639
Db 2629 GAGCAGCAGCCCCCGCTGAGTTCCTGGCAGACAGCCCCCTGCAACAAGTCCCCCTGA 2688
Qy 2640 TGAAGCACCGGGGTGACACGAGAGTACAGAGATGACCTGAGCGGCTCGCGCAATCGCT 2699
Db 2689 TGAAGCACCGGGGTGACACGAGAGTACAGAGATGACCTGAGCGGCTCGCGCAATCGCT 2748
Qy 2700 GGAAGCGCAGGCCCCGCAACCGGGCAACAAATGGTTCAGTGTCTGGCAACCGCAAC 2759
Db 2749 GGAAGCGCAGGCCCCGCAACCGGGCAACAAATGGTTCAGTGTCTGGCAACCGCAAC 2808
Qy 2760 CAGGCTTCTCATGTGTGACTTTCAGCATGCAAGTGAAGATCAAGTTGTCTGAAAACCTGCT 2819
Db 2809 CAGGCTTCTCATGTGTGACTTTCAGCATGCAAGTGAAGATCAAGTTGTCTGAAAACCTGCT 2868
|||||

QY	2820	GAGGAATTTCAAAAACAGCAACGGGTGGCAGAACTGTGGGTGTTCACAAACTTCTG	2879
Db	2869	GAGGAATTTCAAAAACAGCAACGGGTGGCAGAACTGTGGGTGTTCACAAACTTCTG	2928
QY	2880	CTGTCTTCTTCAACAATTCACACCAAGCAATATCCCTGTCCAGCTGCTGTCTGG	2939
Db	2929	CTGTCTTCTTCAACAATTCACACCAAGCAATATCCCTGTCCAGCTGCTGTCTGG	2988
QY	2940	CTACTCGTCAACATCCCTTGTAGTCCGAGAACATCCAGAAAGCTACGTTCAGCT	2999
Db	2989	CTACTCGTCAACATCCCTTGTAGTCCGAGAACATCCAGAAAGCTACGTTCAGCT	3048
QY	3000	GCACCTCAAGCCCAACGCTCTACTCTTCAGGGCGGAAAGCGAGTACAGCTTCGAAAGGTG	3059
Db	3049	GCACCTCAAGCCCAACGCTCTACTCTTCAGGGCGGAAAGCGAGTACAGCTTCGAAAGGTG	3108
QY	3060	GATGGAAGTATCCGCAAGTGCACACAGCTTCGCTCGGACCCGACGTTGAGCCACAA	3119
Db	3109	GATGGAAGTATCCGCAAGTGCACACAGCTTCGCTCGGACCCGACGTTGAGCCACAA	3168
QY	3120	AGAGTCTCTTGTGTAT	3135
Db	3169	AGAGTCTCTTGTGTAT	3184

RESULT 6
AAC98992
ID AAC98992 standard; cDNA, 3094 BP.
XX
AC AAC98992;
XX
DT 09-MAR-2001 (first entry)
XX
DE Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:220.
XX
Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection;
XX
diagnosis; identification; cyrostatic; neuroprotective; nootropic;
XX
immunomodulatory; relaxant; gene therapy; chromosome mapping;
XX
antiinflammatory; cardiant; tissue identification; tissue typing; forensic;
XX
linkage analysis; tissue identification; tissue typing; forensic; neural;
XX
immune system; muscular; reproductive; gastrointestinal; pulmonary;
XX
cardiovascular; renal; proliferative; BB.
XX
OS Homo sapiens.
XX
PN WO20005320-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005989.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM,
XX
WPI, 2000-579444/54.
XX
DR P-PSDB; AAB54227.
XX
PT New nucleic acid that is a pancreatic cancer antigen for preventing,
XX
treating, or ameliorating a medical condition, particularly pancreatic
XX
cancer, or for use in assays for diagnosing a pathological condition.
XX
claim 1; Page 664-665; 1379pp; English.

Query Match	Best Local Similarity	Matches 1441	Conservative	2	Mismatches	0	Indels	1	Gaps	1																																																																														
1592	GTGTTTCAAGAGCAAGTGAACAAAGAGAGCCATGCGGAGACCTGAAAGTCTCAT	1755	ATTCGCGAAATTTTGAACCTTTGCAAAATTTCAATCAATTTTTCGAAAGAAATTGAGCA	1811	122	ATTTCCGAAATTTTGAACCTTTGCAAAATTTCAATCAATTTTTCGAAAGAAATTGAGCA	181	1812	ACGACTTGCCCTGTGGAGAGGCGCCTCAAAATCCAGAAATTACCAAGATTCGG	1872	182	ACGACTTGCCCTGTGGAGAGGCGCCTCAAAATCCAGAAATTACCAAGATTCGG	241	1872	CGATGTCATGCTGAAGAAATTCAGGCGCATGAGCACTGGCGGCTCACTGTGGAAGCA	1931	Db	242	CGATGTCATGCTGAAGAAATTCAGGCGCATGAGCACTGGCGGCTCACTGTGGAAGCA	301	Qy	1932	CAGGAGAGCCCTTGAGAGGCGCCTGAGAAATGGAATTAAGATCCCGGCGGTGAGAAATTT	1991	Db	302	CAGGAGAGCCCTTGAGAGGCGCCTGAGAAATGGAATTAAGATCCCGGCGGTGAGAAATTT	361	Qy	1992	CTGCAAGAGACTTTGAGCTGCGAAGAGGTGTGTTACTCAACCGCTTCTCTCTGCG	2051	Db	362	CTGCAAGAGACTTTGAGCTGCGAAGAGGTGTGTTACTCAACCGCTTCTCTCTGCG	421	Qy	2052	GCCAATGACACCGGCTCAATGCACTAAGACAGAGTCTGAGAGCGGCTGTGCAAAACACACC	2111	Db	422	GCCAATGACACCGGCTCAATGCACTAAGACAGAGTCTGAGAGCGGCTGTGCAAAACACACC	481	Qy	2112	GCCGAGGACGCGGCACTTCAAGGACCTGCGGAGCGCTTTGGCAGAGATCAAGAGATGT	2171	Db	482	GCCGAGGACGCGGCACTTCAAGGACCTGCGGAGCGCTTTGGCAGAGATCAAGAGATGT	541	Qy	2172	GGCAGAGCTCAACGGTACGATGATCAAGATGAGAAATTTCCAGAAAGCTGACGAACTCAA	2231	Db	542	GGCAGAGCTCAACGGTACGATGATCAAGATGAGAAATTTCCAGAAAGCTGACGAACTCAA	601	Qy	2232	GAAAGATTTGATTTGGCAATGACATCTTGTGTTCGCGGAAAGGAGTTCAATCGTCTGG	2291	Db	602	GAAAGATTTGATTTGGCAATGACATCTTGTGTTCGCGGAAAGGAGTTCAATCGTCTGG	661	Qy	2292	CAGCTTCAAGCAAGCTTCTCGGGGAGGGGCTTCAGAGCGCATGTTCTTCTGTTCACAGCA	2351	Db	662	CAGCTTCAAGCAAGCTTCTCGGGGAGGGGCTTCAGAGCGCATGTTCTTCTGTTCACAGCA	721	Qy	2352	CGTCTGCTTATACAGAGCCGGGGGCTGAGCGGCTCAATCAAGTTAAAGTCCAGGAGCA	2411	Db	722	CGTCTGCTTATACAGAGCCGGGGGCTGAGCGGCTCAATCAAGTTAAAGTCCAGGAGCA	781	Qy	2412	GCTCCCGCTCTATGAGTGAAGATTTGAGAGAGCGAAGACGAGTGGGGGTGCCCACTG	2471

```
DB 782 GCTCCGCTCTATGCGATGACGATTGAGAGAGCGAAGACGAGTGGGGTCCCATCAGT 841
QY 2472 CCTGACCTCCGGGGGCGAGGGGAGTCAATCATCGGGGCGGCGAGTTCTCGGTCGAGAT 2531
DB 842 CCTGACCTCCGGGGGCGAGGGGAGTCAATCATCGGGGCGGCGAGTTCTCGGTCGAGAT 901
QY 2532 GAGAGAGTGGGTTGAGAGATCCAGATGGCCATTTGACCTGGCGAGAGAGAGAGAGAGCC 2591
DB 902 GAGAGAGTGGGTTGAGAGATCCAGATGGCCATTTGACCTGGCGAGAGAGAGAGAGAGCC 961
QY 2592 CGCCCTGAGTTCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2651
DB 962 CGCCCTGAGTTCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
QY 2652 GGTGACCGAGAGTCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2711
DB 1022 GGTGACCGAGAGTCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 2712 CCGGACCGCGGACACACATGAGTGCAGTGTGGACCGGACACCGACAGCGTCTCAT 2771
DB 1081 CCGGACCGCGGACACACATGAGTGCAGTGTGGACCGGACACCGACAGCGTCTCAT 1140
QY 2772 GGTGACCTTACGATGCGAGTGAAGATCAGTTGTCTGAAACCTGCTGAGAAATTCAA 2831
DB 1141 GGTGACCTTACGATGCGAGTGAAGATCAGTTGTCTGAAACCTGCTGAGAAATTCAA 1200
QY 2832 AAACGACAGGGGTGGCAGAGAGTGGGTTGGTTCAGAAACCTGCTGAGAAATTCAA 2891
DB 1201 AAACGACAGGGGTGGCAGAGAGTGGGTTGGTTCAGAAACCTGCTGAGAAATTCAA 1260
QY 2892 CAATACACAGAGACATCATCCCTTGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2951
DB 1261 CAATACACAGAGACATCATCCCTTGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 2952 CATCCCTTGTAGTCCGAGAAATCAGTACGTTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3011
DB 1321 CATCCCTTGTAGTCCGAGAAATCAGTACGTTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1380
QY 3012 CACGCTCTACTCTTACGAGCGGAGAAAGCGAGTCAAGTTCGAAAGGTGAGTGAAGTGA 3071
DB 1381 CACGCTCTACTCTTACGAGCGGAGAAAGCGAGTCAAGTTCGAAAGGTGAGTGAAGTGA 1440
QY 3072 CCGGACGTGCACAGCTGTGCTGCGAGCCGACGCTGTGAGGCAAAAGAGTCTTGT 3131
DB 1441 CCGGACGTGCACAGCTGTGCTGCGAGCCGACGCTGTGAGGCAAAAGAGTCTTGT 1500
QY 3132 GTAT 3135
DB 1501 GTAT 1504

RESULT 7
ADQ84807
ID ADQ84807 standard; cDNA, 3997 BP.
AC ADQ84807;
XX
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1621.
XX
XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
XX cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
XX WO2004060270-A2.
XX
XX 22-JUL-2004.
XX
XX 15-OCT-2003; 2003WO-US029126.
XX
XX 18-OCT-2002; 2002US-0418988P.
XX
```

```
XX
XX (GUTH ) GENENTECH INC.
PA (MUND/) MU T D.
PA (ZHOU/) ZHOU Y.
XX
XX Wu TD, Zhou Y;
XX
XX WPI; 2004-534300/51.
XX
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
XX
XX Claim 1; SEQ ID NO 1621; 5504bp; English.
XX
XX
XX The present invention describes an isolated tumour-associated antigenic
XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
XX sequence identity to (a)-(c) or (e) a sequence that hybridises to (a)-
XX (c). Also described: (1) an expression vector comprising the above
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)
XX a process for producing a polypeptide; (4) an isolated polypeptide
XX comprising: (a) an amino acid sequence encoded by any of the above
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-
XX length coding region of the above nucleotide sequences; or (c) a sequence
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)
XX an isolated antibody that binds to the above polypeptide; (7) a process
XX for producing the antibody; (8) an isolated oligopeptide that binds to
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)
XX binding organic molecule that binds to the above polypeptide; (10) a
XX composition of matter comprising the above (chimeric) polypeptide,
XX antibody, oligopeptide or TAT binding organic molecule, in combination
XX with a carrier; (11) an article of manufacture comprising a container and
XX the composition of matter contained within the container; (12) methods of
XX inhibiting the growth of a cell that expresses the above protein, where
XX the growth of the cell is at least in part dependent upon a growth
XX potentiating effect of the above protein; (13) a method of
XX therapeutically treating a mammal having a cancerous tumour comprising
XX cells that express the above protein; (14) a method of determining the
XX presence of a protein in a sample suspected of containing the protein
XX described above; (15) methods of diagnosing the presence of a tumour in a
XX mammal; (16) a method for treating or preventing a cell proliferative
XX disorder associated with increased expression or activity of the above
XX protein; and (17) a method of binding an antibody, oligopeptide or
XX organic molecule to a cell that expresses the protein described above.
XX The TAT sequences have cytostatic activities, and can be used in gene
XX therapy. The composition and methods are useful for diagnosing,
XX preventing or treating cancer. The composition is also used for preparing
XX a medicament for the therapeutic treatment or diagnostic detection of a
XX cell proliferative disorder or cancer. The present sequence represents a
XX human TAT cDNA sequence from the present invention.
XX
XX Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;
XX
XX
XX Query Match 33.1%; Score 1037.4; DB 12; Length 3997;
XX Best Local Similarity 60.2%; Pred. No. 9.5e-267;
XX Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;
XX
XX
XX 21 GCCGACCCGAGATGACGATGAGGGGCGGAGAAATTCGGGATCACTTGAACG 80
DB 150 GAGAGACTCAGAGATGAGCGCTTGGGTGGCCAGACCCCTGGGAGTTAGACACCTTGAGCGC 209
QY 81 TGGAGAGAGCGCCGCCCAACACCTTCAGGAAATCGGTGTCATCAAAATCCAGATGT 140
DB 210 TGGGAGACTCTCTTGGCCAGATGCAAGAGACACTTGACCTCAAGATTAAGCTGCT 269
QY 141 GGATGACACCCGAGAGGAGATTTGAAGTTCCACAAGAGCTCTCTGGAAGTGTCTGGA 200
DB 270 GGAACAACACATGAAATATTGACATTGAGGCTTAATGCAATGCGCAGGTATTCTGAC 329
QY 201 TGCAGTTGCAACCACTCAACTGCTGGAAGGTGACTATTGCTGAGTTCCTGA 260
XX
```

Dh 330 ACAAGTGGAGGCGTTAAACCTGTAGATGTGACTTGGGATGAGTTTCAAA 389
Qy 261 TCACAAAAGATCACGGTGTGGCTGGATCTCTTAAACCCATTGTGAAACAGATTAGAG 320
Db 390 TACTCATCTCTAATCTGGAATTTGGCTTGAACCTATGAAACCAATCTATTGGCAAAATACAG 449
Qy 321 GCCAAACACGTTTGTGTAAAGTTGTGTGAATTTCTTCGCGCTAGACACACACAAC 380
Db 450 GCCAAAGATGTGTGTCTTGGCTTACCTTAAATTTTTCACCTTATCTGTGTAGCT 509
Qy 381 CCAAGAAATCTCAAGAGTACTGTTCGCGCTGACAGTGAAGCAGGACTTGGCTCAGG 440
Db 510 ACAAGAAAGATATCAAGATATCTTGTGCTTGCATTTAAGAGACATCTGTGAAGA 569
Qy 441 CAGGTTGAGCTGTATGACACGACGCACTCTTGTGATTTCACATTTGTGCAATGTGA 500
Db 570 GCGTTTGACCTGTGCTACACACAGCGGCTTCTACGCTCCATCTCTGCGAGTGGGA 629
Qy 501 GATTGGGGAATTTGATGAAGCCTTGAACAGAGCCTTAGCAAAAATTAATACATACC 560
Db 630 AATAGAGATTAAGATGAAACGCTGACCGAGACCTTCAAGGTGAACGATATTTTCC 689
Qy 561 TCAGCAGACGCACTAGAGGACAAATCTGTGAATTTTCACTAACCAATTGACAAAC 620
Db 690 TGGCAGACGACACTGCTTGAAGATATCTAGAAATTCATCAGAAAGCAGTGGCCAGAC 749
Qy 621 ACCAGCAGATCAATTTTCCAGCTCTTACAGATTTGCCGTGCTAGAGATGTATGGAAT 680
Db 750 ACCGTGAGTGGATTTTCCAGGTGCTCGAAATTTCTCGAAAGTTGAAATGTACGCGCAT 809
Qy 681 CCGGTTGACACCGGCGCAAGGACAGGGAAGGACAGAGATCAATCTGCGCTGCAACAC 740
Db 810 CAATTTTCAATAGCTTCTGACAGGAAAGAACCAAAATTTCACTGACAGTTTCCCAT 869
Qy 741 GGGAAATCTAGTGTTCAGGGTTTCACTAATGATCAATGCTTCAATCTGCGCAAGGTGC 800
Db 870 GGGTGTACTGTGTTCAGGGCACCAAAATTCAAACATTTTCACTGTGCTCAAGGTCCG 929
Qy 801 GAAGCTGAGCTTCAGAGGAAAGGCTTTCATCACTAGCTCCGCGCAGATGCCAATAGTGC 860
Db 930 TAAACTAAGCTTCAGAGGAAAGATTTCTTATCAATTCATCAAGAGTTTCAATGACC 989
Qy 861 GTACAGGATACCTTGAATTTCTGATGCGCAGTCCGAGATTTTGCAGATCTTCTGAA 920
Db 990 TTACAGGACATATTAAGATTTTGTGGTAGTAGATGATGTATGAACCTTCTGAA 1049
Qy 921 AATCTGTGTGAACATGATGCTTCTTATGACTTTTGAAGAGCCCAACCAAGCCCA 980
Db 1050 GATTTGTGTGAGATCACACCTTTTATGACTTTTGAACCAACCTTAAGCCAAAGCAA 1109
Qy 981 GCCCGTCTCTTAAAGCGGGGTCAATTTCCGTTCACTGTGCTGCACTCAAGACAGGT 1040
Db 1110 AGCGCTCTTCTTACGCGGGGCTCTCTCTTCAATACAGTGAAGAACTCAGAAACAAT 1169
Qy 1041 TCTGACTATGTTAAGAGAGAGACATAAGAGGTGAGTTGAAAGAAAGACAGCAA 1100
Db 1170 AGTAGATTTATTTCAAGACAGTGAATGAAGAGATTTCAATAGAAAGAGCAGCAA 1229
Qy 1101 GATTCAATTTATCTCGAGCCTTGTTCACAGCCTTACAGAACTGAAATTCGAAAGTGTGA 1160
Db 1230 GACCCAACGTCC-----GTTTCAAGCTCTGACTGACCAACCTAACCAA 1271
Qy 1161 GCGGTCTCAGACAGACACAGCCTTACATTTGAGAAAGTCCGGAATCTCAAGGGGCCA 1220
Db 1272 ACAGAGCATCTATTTCCCAAGGAGATTGAGACTCTGCTCCCAATCTTCAAGCATGC 1331
Qy 1221 GAGCTGCGGCGAGAAAGAACGAAAGTTTCCGCGGGGAGCGGGGTGCGACCCGAG 1280
Db 1332 CTTTATCTGCTCTCTCTCTCACTGTGTCTCTCTGCTCTGCGCTGCAAGTTTATGACAG 1391
Qy 1281 CCCTGCGCCGAGAGAGCCCGCGGGTAAACAGCAGCGGACGAGCCCTCGCGCC 1340
Db 1392 CAGCAGCTCCTCAACATCCCAAGTTTCTTCTAGTCAAGATCTCAGACTGCAAGAGCG 1451

Qy 1341 CAGGAGGAAGAGAGAGG---TGTTAAGATAGAGACCCAGACAGTAACTCAGCC 1397
Db 1452 CAGTGAAGCATGTGGCTGAGAGGCCCCCAACACCATGGGCCCTCGGGCCCCCGC 1511
Qy 1398 CCGCAGCCAGACAGCAGGCTCCTGACTGGCAGTCTCACTTTCGAGCTGTCTGTGA 1457
Db 1512 ACTCAGACCTGTGTCAAGGCTTTCACAGAAAGATCTCAGCTTCTCCCTCCAGCGGAA 1571
Qy 1458 CTGCGAGGGGGAGTGGCCCTTCCAACTGACTGTCTCCCACTGAGCCCGACAC 1517
Db 1572 GAGCCCTGAGCTGAGCCCTGACTATTCAGGTGCTT-----TGGGCCCACTGA 1622
Qy 1518 CAGCAGGCTCTCCCTTGAATCAGCCCGCTGTGAAATGACCAAGCCCTGCCCGGACGA 1577
Db 1623 ACAGGGCTATCCCACTCTGAGCCCTGTCTCAGTGAATGTGGCGAGCCGGAGTGA 1682
Qy 1578 CGATGAGGATGAGGGGCGGAGAGAGATTTCCCACTGATTAAGCGTACTTCAATAGCTAA 1637
Db 1683 CTGCGAG---GAGCCAGACAGAGCGCGTGCCTGACAGAGAGCCCTACTTATAGTCA 1739
Qy 1638 GGAAGTGTCTACCAAGGCGAAATATGTGAAGATCTCGAAATTCATCTTGTGTT 1697
Db 1740 AGAGATTTCTGCTACAGAAAGAACTATCTCAGAGATTTAAGATTTATTCGCTGTGTT 1799
Qy 1698 TCAGACGACGTGAGCAAGAGAGAGCCCAATGCGGAGACACTGAAAGTCTCATATTCCC 1757
Db 1800 CCGCAGCGCATGTGTGAAGAGAGCCCATGCTGTGACCTGTATGACGCTGCTTCTC 1859
Qy 1758 GAATTTTGAACCTTTCACAAATTTTCACTAATTTTCTCAAGAAATTTGAGCAAGACT 1817
Db 1860 CAACATCGATCCCATCTATAGATTTCCACAGAGCTTCTGCGGAGTGAAGAGGCT 1919
Qy 1818 TGCCTGTGGAAAGGCGCTCAAAATGCCAAATCAGAG---ATTACAAAGAAATCGGCGA 1874
Db 1920 GGCACCTGTGGAAAGGCGCTCCCAAGCCCAACAAAGGCAATCAACGAATCCGGGGA 1979
Qy 1875 TGTCACTGTAAGAAATTTCAAGGCGATGAAAGCACTGGGGGCTCACTGTGGAAGACAG 1934
Db 1980 CATCTGCTCAGAAACATGTGGCCAGTTAAGAGTTTACAGCTTCTCCAAAGACATGA 2039
Qy 1935 CGAGGCTTGGAGGCGCTGAGAGATGAATCAAGACTCCGCGGCTGAGAACTTCTG 1994
Db 2040 CGAGGTCTTAACGAATGGAAGAGCTACCAACGCTGTAAAGAAATTTGAGGCAATGTA 2099
Qy 1995 CAGAGACTTGAAGCTCAGAGAGGTGTTCCTTACCGCTCAGACACTTCTCTCGGCGC 2054
Db 2100 CAAGGATTTGAGCTCAGAGAGGTCTGTACTTGCCTTCAACACGTTCTCTGAAAGCC 2159
Qy 2055 ACTGCAACCGGCTATGCACTACAGAGAGTCTGAGAGCGGCTGTGGAACACACCCGCC 2114
Db 2160 CATCCAGCGGCTGTCACTACCGCTGTGCTGCGCCGCTATATGCGAATTAACGCC 2219
Qy 2115 GAGCCAGCGCACTTACAGGAGCTGCCAGCGGCTTGGAGAGATCAAGAGATGTGTGC 2174
Db 2220 CGGGACCATGACTAGCTGATGCCATGACGCTTGAAGCAATACAGAGGTGACAC 2279
Qy 2175 ACAGTCCACGATCATGATTAAGATGAAGATTTTCAAGAGCTCAGAACTCAAGAA 2234
Db 2280 CACACTACAGCACTTCTATCCGCTGGAAGAACTGCAAGAACTTACAGAGCTGACGCG 2339
Qy 2235 AGATTTGATTTGAGATTAAGATTTGTGTTCGCGGAAGGAGTTATCTGCTGCGGAG 2294
Db 2340 GAGCTGTGTGTGATGAAGAACTCATTTCTCTGCAAGGAGTTATCTGAGGGGTG 2399
Qy 2295 CCTCAGCAAGCTCTCGGGAAAGGGGCTCAGACAGCAGATGTTCTTCTGTTCAACAGCT 2354
Db 2400 CTTTCAAGAGCTACCAAGAGGGCTTGCAGAGAGAGATTTTCTGTCTCAGATAT 2459
Qy 2355 CTTGCTATACAGAGCGGGGCTGACGCTTCGAATCAAGTTTAAAGTCAAGGGCAGCT 2414
Db 2460 GTTGTGTATACAAAGCAAGAGGTTGCAAGGACCAACCTTCCGAGTCCGGGGCTCTCT 2519

QY	2415	CCCGCTCTAATGGCAATGAAGATTGAGAGAGAGGAAAGACAGATGGGGGGTGGCCCACTGGCT	2474
Db	2520	TCCCTCTCCAAAGGCATGCTGCTGTGGAAAGAAATGATACAGATGTGTCTGTTCCACACTGTTT	2579
QY	2475	GACCTTCGGGGCCAGCGGCAGTCCATCATGTGTGGCCCGCAGTTCTTGTGTCCAGATGGA	2534
Db	2580	CACCATCTACCCGGCTCAGAAAAACATGTGTGTGGCAGCCAGCACTCGCTGGAGAAAGA	2639
QY	2535	GAAGTGGGTTTAGGACATCCAGATGGCAATTGACCTGGCGGAGAAAGACAGACGCCCGC	2594
Db	2640	GAAGTGATGCTGTGAACCTGAACTCCGCCATCCAGCACCAAGAGTGGCGGTACACGGC	2699
QY	2595	CCCTGAAGTTCCTGGCCAGCA-----GCCCCCTGACCAACAAGTCCCTGATGAAGCCAC	2648
Db	2700	CCCTGCACTGGCCAGGCGCGCATGTGTGTGACATCTGTCCCCCGAGATCCCCCAACGAGATTC	2759
QY	2649	CGCGGCTACCAAGAGATCAGAGATGACTAGCGCCCTCGGCGCATGTGCTGGAGCCGCA	2708
Db	2760	TCTG---GAGAGAGAGTCAAGAAATGAATGCTCGGGGTGTCTCGCAGCTCTCTGGAGGGGCA	2816
QY	2709	GGCCCGCACCGCGGCAACAATAATGATGCAGTGTGCTGGCAGCCGCAACACAGCGTCTC	2768
Db	2817	TGGCCATGACCGGGGCAACACCAATGCAAGTGTGTGCTGTATCCGGAACACACAGCGTCT	2876
QY	2769	CATGTGACATTACGACATCGCAGTGGAGAAATCATGTTGTTGGAAACTGCTGAGGAAATT	2828
Db	2877	CAGGCGAATCCACAGTGCAGCTGTCTGAGAACAGAGCTTTCAGGAATATCTGTAGAAAGTT	2936
QY	2829	CAAAAAACGCAACGGGTGGAGAAAGCTGTGGGTGTGTTCAAAACTTCTGCTGTTCTT	2888
Db	2937	CAAAAACAGTATGCTGTGCAAGAAAGCTCTGGGTGTCTTTTACAACTTCTGTGTTCTT	2996
QY	2889	CTACAAATCACACCAAGACAAATCATCTCCCTTTCAGACCTGCTCTGTCTGAGCTACGCT	2948
Db	2997	CTACAAAACTCATCAGATGATCTACCCCACTGGCCAGCTCTCCGCTGCTGGGCTACAGCGT	3056
QY	2949	CACCATCCCTCTGAGTCCGAGAAACATCCAGAAAGACTAGCTTTCAAGCTGCACCTCA	3008
Db	3057	GAGCATCCCGCAGGAGGCGCCGATGGCATACAAAGACTATGTATTTTCAAGCTCCAGTTCAA	3116
QY	3009	GTCCACAGCTCTACTACTTTCAGGGCGGAAAGCGAGTACAGTTTCGAAAGGTGATGAAAT	3068
Db	3117	ATCCACGCTCACTACTTCTTCGGGGCTGAGAGACATGACATTTTGAAGGTGATGAGGT	3176
QY	3069	GATCCGACGTGCCACACAGCTCTGCTCGGCACCCC 3103	
Db	3177	GATCCAGGGGCCACGACAGCTCAGCCGGAGGGGCC 3211	
RESULT 8			
ADR25675			
ID	ADR25675 standard; DNA; 3997 BP.		
XX	ADR25675;		
AC	ADR25675;		
DT	21-OCT-2004 (first entry)		
XX	Breast cancer prognosis marker #1536.		
DE	Breast cancer; prognosis; gene expression; diagnosis.		
XX	de; breast cancer; prognosis; gene expression; diagnosis.		
KW	Homo sapiens.		
OS	Homo sapiens.		
XX	MO2004065545-A2.		
PN	05-AUG-2004.		
XX	15-JAN-2004; 2004WO-US001100.		
PF	15-JAN-2003; 2003US-00342887.		
XX	(ROSE-) ROSETTA INPHARMATICS LLC.		
XX	(NECA-) NETHERLANDS CANCER INST.		
PA	(ROSE-) ROSETTA INPHARMATICS LLC.		

PI	Van't Veer LJ, He Y;
XX	WPL; 2004-593473/57.
XX	
PT	Classifying a breast cancer patient according to prognosis comprises
PT	determining the similarity between the level of expression of each of
PT	five genes in a cell sample taken from patient, to control levels.
XX	
PS	Disclosure; SEQ ID NO 1536; 226bp; English.
XX	
CC	The invention relates to a method of classifying a breast cancer patient
CC	according to prognosis by determining the similarity between the level of
CC	expression of each of five genes for which markers are listed in the
CC	specification, in a cell sample taken from the breast cancer patient, to
CC	control levels of expression for each respective five genes to obtain a
CC	patient similarity value. The methods are useful for classifying a breast
CC	cancer patient according to prognosis. Kits and computer program products
CC	are useful for data analysis using the diagnostic, prognostic and
CC	statistical methods of the invention. This sequence corresponds to a
CC	marker used in the method of the invention.
XX	
SQ	Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;
	Query Match 33.1%; Score 1037.4; DB 13; Length 3997;
	Best Local Similarity 60.2%; Pred. No. 9.5e-267;
	Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7.
OY	21 GCCGACCCCGAGATCAAGACTGGGGGCCCGCGAATAATTCGGGATCAGTACCCTTGGAACG 80
DB	150 GCAGATCTGCAGGATGGCGCTTGGGTCCCAAGCCCTGTGGAGTTAGCACCTTGAACC 209
OY	81 TGACAGAAGCCGCCCAACAACCTTCAGAAAATCGTGTCAATCAAATTCAGATCT 140
DB	210 TGGGCACACTCTCTTGCCCAAGAAAGCAAGAACACACTGCACCTCAGAGTAAGCTCT 269
OY	141 GGATGACACCCAGGAGCATTTGAAGTTCACAAAGAGCTCTGGAAAGTGGCTGCGA 200
DB	270 GGACAAACCAATGSAATATTTCATTAGCTTAATGCGATGCCAGATTATTCAGC 329
OY	201 TGCAGTTTGCAACCACTCAACCTCTGAGAGGTGACTATTTTGGCCTCGAGTTTCTGA 260
DB	330 ACAAGTGTGGAAGCTTTAAACCTGTAGATGTGACTACTTCGGAGTAGAGTTTCAAAA 389
OY	261 TCACAAAAAAGATCAAGSTGTGGCTGATCTCTTAAACCATTGTGAAACAGATTAGAG 320
DB	390 TACTCAGTCTTACGTGATTGGCTTGAACCTATGAACCCATCATTTAGGCAAAATAGAG 449
OY	321 GCCAAACACGTTTGTGTAAGTTTGTGTGAATTCCTTCCGCTGCACACACAACACT 380
DB	450 GCCAAAGAAATGTGCTTGGCCTTACGTATTAATTTTTCCACTGATCTCTGCTGACGT 509
OY	381 CCAGAAAGAACTCAAGATGACTCTGTTGCGCTGCAGGTGAAGCAGACTTGGCTCAAGG 440
DB	510 ACAAGGAAGATATTCAGATGACTGTTTGGCTTCCAATTAAAGAGAGCTGCTGGAAGA 569
OY	441 CAGGTTACGCTGTAATGACACCAAGGAGCTCTTGATTTTCAACATTTGCAATTCGA 500
DB	570 GCGTTGACCTGTCTGACCAACAGGGGCCCTTCTCAGCTGCCATCTCTCTGCACTGGA 629
OY	501 GATTGGGATTTTGTATGAAGCTTTGACAGAGCACTTAGCAAAAATAATATACATAC 560
DB	630 AATAGGAGATTAAGATGAAGCGTGAACCGAGAGCACTCAAAGTGAACGATATTTGCC 689
OY	561 TCAGCAGACGCACTAAGAGCAAAATCTGTGAATTTTCAACATACCATTTGACAAAC 620
DB	690 TTGGCAGACGACCTGCTTGAAGAGATCTAAGATTCATCAAGAAACGTTGGGCAAGC 749
OY	621 ACCAGCAGAAATCAGATTTTCAGCTCTAGAGATTGCGGTGGCTAGAGATGATGAAAT 680
DB	750 ACCTGCTGAGTCTGATTTCCAGAGTCTCTGAAATTTGCTGAAAGTTTGAATATGACGAT 809
OY	681 CCGGTGACCCCGGCAAGAGCAGGGAAGGCAAGAGATCAATCTGCGCTTGGCCACAC 740

Db	810	CAGATTTCACATGCGTCTTGACAGGGAAGAAACCAAGATTCACATGCGAGTTTCCCAAT	869
Qy	741	GGGAATTCCTAGCTGTTTCAGGGTTTTCATCATAAGTCAATGCCTTTCACATGCGGCAAGAGTCG	800
Db	870	GGGTGTACTCGGTGTTCCAGGGCACCAACCAATCAACCTTTCACATGCTGTCCAGAGTCG	929
Qy	801	GAAGCTGAGCTTCAAGAGAAAGCGTTTCTCATCAAGCTCCGGCCAGATGCCAATAGTGC	860
Db	930	TAAACTAAGCTTCAAGAGAAAGATTTCTTATCAAACTTCAATCCAGAGGTTCAATGAC	989
Qy	861	GTACCAAGATACCTTGGAAATTCGTGATGGCCAGTCGGGATTTCTGCAAGTCTTCTGGAA	920
Db	990	TTACCAAGACACATTAGAAATTTGTTGGTAGTAGAATGATGTAAGAACTTCTGGAA	1049
Qy	921	AATCTGTGTGAACATCAATGCGCTTCTTGAACCTTTTGAAGGCCCAACCAAGCCCA	980
Db	1050	GATTGTGTGAGATATCACACTTTTATGACCTTTTGAACCACTTAAGCCAAAGCAAA	11089
Qy	981	GCCCGCTCTTTCAGCCGGGGGATCATATTCGTTTCAGTGTGCGACTCAGAAAGAGT	1040
Db	1110	AGCGGTCTTCTTACCGCGGGGCTCTCTTCAAGATACGTGGAAAGAACTCAGAAACACT	1169
Qy	1041	TCTGACATATGTTAAAGAGAGACATTAAGAGGTGCAATTTGAAGAAAGCACAGCA	11000
Db	1170	AGTAGATTATTCAAGACAGCTGGAATGAGAAATTCATATGAAAGAAAGCACAGCA	1229
Qy	1101	GATTCTATTCTATCCGGAGCTTGTGTTCAACGCTACAGACTGAAATTCGGAAGTCTGGA	1166
Db	1230	GACCCACACGTC-----GTTGAGCTCTGACTGACACTACAA	1271
Qy	1161	GCACTCTACAGACAGACACAGCTTACATTTGAGAAAGTGCCTGATCTCCAGGGGAGCA	1220
Db	1272	ACAAAGACATCATTTCCCGAGGAGATTAGAGACTCTGCTCCCATTTTCAAGCAATGC	1331
Qy	1221	GAGCTGCCGGCGAGGAAGAAACCGAAGTTTCCGCGCGGAGACCGGGGTGCAACCGAG	1280
Db	1332	CTTTTACTCGCTCTCTCCCTCACACTCTGCTCCCTCTGAGCTGACAGTTTAAAGACAG	1391
Qy	1281	CCCTGCGCGAGAGGAAGCCCCGGGGGTAAACAAGAGCGGACGGAAGCGGCTCGGCGCC	1340
Db	1392	CAGCAGCTCCCTCAAGATATCCCAAGTTTCTTACGTCAGAGTCCAGCTGACAGAGAGCG	1451
Qy	1341	CACGAGGAAGAGAGAGAG--TCGTAAAGATAGAACCCAGCAGATTAACCTCAGCC	1397
Db	1452	CAGTGGAGCAATGGCTGGAGGCCCCGACACACATCGGCCAGAGCCCTTCGGGCCCCCGC	1511
Qy	1398	CCCGCAGCCAAAGACAGAGCTCCCTGATCTGGCAGTCTTCACCTTTCGAGCTGTGTGAA	1457
Db	1512	ACTCAGGCTGTGTCAAGGCTTTTCCAAGAAAGTCTCAGGCTTCTCCCTCAAGCCGGA	1571
Qy	1458	CTCGCAGGGGGAGTGGCCCTCTGCAAGGTACCTTGTCTCCCACTGAGCCCGGACAC	1517
Db	1572	GAGCCCCCTGAGTGTGAGCCCTGTGATTCAGGTGCTT-----TGGGCCCAAGCTGA	1622
Qy	1518	CAACGAGGCTCTCTCCCTTGATCAGCGCGCTGTAATGACAGGCGTCCCGCGAGCGGA	1577
Db	1623	ACAGGGCTCATCTCCCACTCTGAGCCCTGTCTCTAGTATGCTGCGGAGCGGGATGGA	1682
Qy	1578	CGATGAGAGTAGAGGCGCGAGGAAGATTTCCAACTGATTAAGGCTTCTCATAGCTTA	1637
Db	1683	CTGCGAG--GAGCCAGACACAAAGCGGTGCTGTGACAGAGGCTTCTCATATGTCAA	1739
Qy	1638	GGAAGTGTCTTACCAACGAGCGAAATATCTGAAGATCTCGAAGTATCACTTGTGTGT	1697
Db	1740	AGAAATTCCTGCTTCAAGAACAACTATCTCAAGATTTTAAGATTAATTAACCGTGTGT	1799
Qy	1698	TCAGAGCACAATGAGCAAAAGAGACGCGCATGCGGGAACACTGAAGATCTCATATTTCC	1757
Db	1800	CCGCAAGCGCAATGGTGAAGAGAGACGCAATGCTGTGCACTGTGATGACGCTGTCTTC	1859
Qy	1758	GAATTTTGAACCTTTGACAAATTTCAATCTAATTTTCTCAAGAAATGAGACGACT	1817

Db	1860	CAACATGATCCATTCTATGATGTTCCACAGAGGCTTCTGCGGAGGTGACGAGAGGCT	1919
OY	1818	TGCCCTGTGGGAAAGCGGCTCAATGCGCCAAATTCAGAG---ATTACCAAGAAATCGGCGCA	1874
Db	1920	GGCACTCTGGGAAAGGCGCTTCCAAAGCCCAACAAAGGACAGTCATCAAGAAATTCGGGGA	1979
OY	1875	TGTCATGTGTGAAGAACATTTCAGGGCATGAAGCACCTGGCGGCTCACCTGTGGAAGCAAG	1934
Db	1980	CATCCTGTCTCAGAGAACATGCGCCAGTTTAAAGGAGTTTACACAGTACTCTTCAAAAGACATGA	2039
OY	1935	CGAAGGCTTGGAGGCGCTTGAGAGAAATGGAATCAAGAGTCCCGGCGGTGAGAACTTCTG	1994
Db	2040	CGAGGTCTCTTAACAACAATCTGAGAAAAGGTATCCAAACGCTGTAAAGAAAGTTGAGAGCAGTGA	2099
OY	1995	CAGAGACTTTGAGTGTGAGAAAGGTGTATCTACACGCTCAACACTTCTCTCTGCGGCGC	2054
Db	2100	CAAGAGATTGTAAGTGTGAGAAAGGTCTCTACTTGGCTCTCAACACGTTCTGTGTGAAGCC	2159
OY	2055	ACTGCACCGGCTCATGCACTTACAAAGCAGGTCCTTGAGACGGGCTGTGCAACACACCCGCC	2114
Db	2160	CATCCAGCGGCTGTGCATTAACGCGCTGTGCTGCGCGCCGCTTATGCGGACATTACAGGCC	2219
OY	2115	GAGCCAGCGGCACCTTCAGGAGACTGTCCGACGCCGCTTTGGCAGAAATCACCGAAGATGTGGC	2174
Db	2220	CGGGCACATGATCACTACGCTGACCTGCACATGACGCCCTGAAAGCCATCAACAAGGTGACAC	2279
OY	2175	ACAGCTCCACGGGTACGATGATCAAGATGAGAAATTTCCAGAGGCTGACAGACCTCAAGAA	2234
Db	2280	CACACTACAGCAATTTCTCATTCGGGCTGGAGAACCTGCAAGAGTAAACGAGAGCTGCAGCG	2339
OY	2235	AGATTGTATTTGGCATTTGATGACATCTTGTGTTCCGGGAAAGGAGTTCAATCCGCTGGGAG	2294
Db	2340	GGACCTGGTGGGACATAAGAAACGTCAATGCTCCTGGGAGGAGTTATCTGTAGGGGCTG	2399
OY	2295	CCTACAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAGCGCATTTCTTCTGTTCACAGACGT	2354
Db	2400	CTTTCAACAAGTCTCAACAAGAGGGCTCGACAGCAGAGATTTTTTTCTGTTCTCAATATAT	2459
OY	2355	CTGTCTTATPACACGAGCCGGGGGCGTACGCGGCTCCAACTGATTAAAGTCCACGGGACGT	2414
Db	2460	GTTCCTGTACACAGCAAAAGGATTTGAGGAGACAGGCATTTCCGGATTCGGGGCTCTCT	2519
OY	2415	CCCGCTCTATGGCATGACATTGAGAGAGCAGAACGAGTGGGGGTGCCCATCTGCTCT	2474
Db	2520	TCCCTCCAAAGGACATGTGTGGAGAAAGTGAATTAAGATGTGTCTTCAACACTGTTT	2579
OY	2475	GACCTTCGGGGGCGACGCGCATCTCATCTGTGGCGCCGACATTTCTGTGTCGAGATGGA	2534
Db	2580	CACCATCTAACCGGCTCAAGAAAACATCTGTGTGGCAGCCAGCACTCGCTGAGAAAGA	2639
OY	2535	GAACTGGGTGAGGACATTCAGATGTGCAATGTGCATGTGGCGGAAAGAGCAGCAGCCCCGC	2594
Db	2640	GAACTGTGATGTGACCTGAACTTCGCGGATTCAGACAGCCAAAGATGGCGGTGACACGGC	2699
OY	2595	CCCTGAGATCTCTGCGCAGCA-----GCCGCCCTGACACAACTATCCCTGATGAACACAC	2648
Db	2700	CCCTGCACTGCGAGGCGGCACTGTGTGCACTGTCTCCCCCAAGTCCCCCAAGAGGTATC	2759
OY	2649	CGCGGCTGACCAAGAGTCAAGAGATGACTTGAGCGCTCGCGCACTATCGTGGAGCGCA	2708
Db	2760	TCTG---GAGCAGGAGTCAAGAAAGTGAATGTCTGGGGGTGTCTCGCAGCTCCCTGAGGGGCGCA	2816
OY	2709	GGCCCGGACCGGCGGCAACCAATGTGTGACGTGTGTGGCAACGGAACACAGCGCTTC	2768
Db	2817	TGGCACAACCGGGCCAAACACCAATGACGTTGTGTGTATACCGGAACACAGCGGTTC	2876
OY	2769	CATGTGTGAATTCAGCATTCGACATGTGAGATCAGTTGTCTGGAACCTGTGCTGAGAAAT	2828
Db	2877	CAGGGCAGACCAAGTGCACAGTCTTCAGAAACAGCTTTTACAGATATCTGTGAAGAAAGTT	2936
OY	2829	CAAAAACAGCACGGGTGCGAGAAAGCTGTGGGTGTGTTCACAACTTCTGCTGTTCTT	2888
Db	2937	CAAAAACAGTCATGTGCTGGCAGAAAGCTGTGGGTGTGTTCACAACTTCTGTTGTTCTT	2996

QY 2889 CTACAAATCACACAGACATATCCCTTGGCAGCCTGCTCTGCTGCTACTGCT 2948
DB 2997 CTACAAAATCTATCTAGAGATGATCACTACCACTGGCCAGCCTCCGCTGCTGGCTACAGCCT 3056
QY 2949 CACCATCCCTCTGTAGTCCGAGAACATCCAGAAAGCTAGTGTTCAGCTGCACTTCA 3008
DB 3057 GAGCATCCCGAGGAGCCGATGCGATACAAAGACTATGTTTTCAGCTCCAGTTCA 3116
QY 3009 GTCCACGCTTACTCTTCAAGGCGGAAAGCAGTACAGTTGAAAGGTGATGGAAGT 3068
DB 3117 ATCCCACTCTACTCTTCCGAGCTGAGCAAGTACATTTGAAAGGTGATGGAAGT 3176
QY 3069 GATCCGACGTGCGACCAAGCTTGCCTGCGGACCC 3103
DB 3177 GATCCAGGGGGCCAGAGCTCAGCCGGGAGGGGCC 3211
RESULT 9
AD564452
ID AD564452 strand: DNA; 3997 BP.
AC AD564452;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human KIAA0793 gene.
XX
KM Human; interferon; IFN; gene; ds; KIAA0793.
XX
OS Homo sapiens.
XX
PN US2004185489-A1.
XX
PD 23-SEP-2004.
XX
PF 17-MAR-2004; 2004US-00802432.
XX
PR 21-MAR-2003; 2003EP-0006263.
XX
PA (CERT/) CERTA U.
XX (FOSE/) FOSER S.
XX PA (MEYE/) MEYER K.
XX
PI CertA U, Foser S, Weyer K;
XX
DR WPI: 2004-689187/67.
XX DR GENBANK; AB018336.
XX
PT Determination of biological activity of compound which can modulate gene
PT transcription, involves contacting host with compound, determining
PT general transcriptional gene response of host, and quantitating gene
PT response induced by compound.
XX
PS Claim 8; SEQ ID NO 16; 53bp; English.
XX
CC The invention relates to a method for determining the biological activity
CC of a compound which can modulate gene transcription. The method involves
CC contacting a host with a compound, determining the general
CC transcriptional gene response of the host and quantitating the gene
CC response induced by the compound. This gene transcription assay method is
CC useful for the determination of the biological activity of a compound
CC which can modulate gene transcription. The present sequence is the human
CC KIAA0793 gene. This sequence is used to estimate the influence of the
CC pegylation site on the transcriptional activity of interferon (IFN) by
CC measuring this IFN-induced gene expression pattern of the melanoma cell
CC line MB15 via the oligonucleotide array technology.
XX
SQ Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;

Query Match 33.1%; Score 1037.4; DB 13; Length 3997;
Best Local Similarity 60.2%; Pred. No. 9.5e-267;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

QY 21 GCCGACCCAGGATTCAGATG3666CCCCGGAATAATCGGGGATCAGTACCTTGGAAAG 80
DB 150 GCAGACTGCAAGGATGCGCTTGGGTGCCAGACCCCTGTGGGAGTTAGACACCTTGAGCC 209
QY 81 TGGACAGAAAGCCGCCCAACACCTTCAGAGAAATCGTGTTCATCAAAATTCAGATGCT 140
DB 210 TGGGAGAGCTCTTGGCCAGATGCAAGAGAGAGACCTGCACTCAAGATTAACCTGCT 269
QY 141 GGAATGACCCAGAGAGGATTTGAAGTCCAAAGAGCTCCTGGGAAGGTGCTGTGGA 200
DB 270 GAGCAACACCATGGAATATTTGACATTGAGCTTAAATGCGATGCGCAGGTATTTACTGAC 329
QY 201 TGCAGTTGCAACCACTCAACCTCGTGAAGGTGATCATATTTTGGCTGAGTTTCTGA 260
DB 330 ACAAGTGTGGAAGGCTTTAAACCTGTGAAATGTGATCTACTCTGGGATGAGGTTCAAAA 389
QY 261 TCACAAAAAGATCACGGTGTGCTGGAATCTCTAAACCATTTGAAAACAGATTGAAG 320
DB 390 TACTCAGTCTACTGGAATTTGGCTTGAACCTATGAACCCATCATTAAGCAAAATACGAAG 449
QY 321 GCCAAGCAGCTGTGTGTTAAGTTGTGTAATTTCTTCGCGCTGACACACAACT 380
DB 450 GCAAGAGATGTGTGCTTCCGCTAGCTGTAAATTTTTCACCTGATCTGCTGACGCT 509
QY 381 CCAGAGAACTCAACAGATACCTGTCGCGCTGCGAGTGAAGCAGAGCTTGGCTCAAAG 440
DB 510 ACAAGAAATATATACAAATACCTGTTGCCCTGCACTTAAAGAGACCTGCTGGAAGA 569
QY 441 CAGGTGACGTGTATGACACCAAGCGAGCTCTCTGATTTCAACATTTGCAATTCGA 500
DB 570 GCGTTGACCTGTGCTGACACCAAGCGCCCTTCTCAAGTCCATCTCCGCACTGGA 629
QY 501 GATTGGGATTTTGAAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 560
DB 630 AATTAGAGATTTCAATGAAG 689
QY 561 TCAGCAAGACGACCTAGAGACAAATATGTTGAATTTTCAACATTAACCATTTGAACAA 620
DB 690 TGGCAGAGAGACCTGCTTGAAGAAATCTAGAAATTCATCAGAGAGAGAGAGAGAGAG 749
QY 621 ACCAGCAATATCAATTTCCAGCTCTTGAAGATTTGCCGTGAGAGATTTATGAT 680
DB 750 ACCGTGCTAGTGGATTCAGAGTCTGAAATTTGCAAGATTTGAAATTTGAGGCAAT 809
QY 681 CCGGTTGACCGGCGCAAG 740
DB 810 CAGATTTCAATGCTTCTGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 869
QY 741 GGGAAATCTAGTGTTCAGGGTTTCAATGATCAATGCTTCAATCTGAGGCGCAAGGTGCG 800
DB 870 GGGTATCTCGTGTTCAGGGGACACCAAAATCAACCTTCACTGATGATGATGATGAT 929
QY 801 GAGGTGAGCTTCAAG 860
DB 930 TAAACTTAAGCTTCAAG 989
QY 861 GACAGAGATCTTGAATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 920
DB 990 TTACAGAGACATTAAGATTTTGTGGTATGATGATGATGATGATGATGATGATGATG 1049
QY 921 AATCTGTGTAACATCAATGCTTCTTGAATTTTGAAGAGAGAGAGAGAGAGAGAGAGAG 980
DB 1050 GATTTGTGAGAGATTCACACTTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 1109
QY 961 GCGGCTCTCTTGAAG 1040
DB 1110 AGCGCTCTCTTGAAG 1169
QY 1041 TCTGACTATGTTAAAG 1100
DB 1170 AGTAGATTTTCAAG 1229

QY	1101	AATTTCATTCTATTCGGAGACCTTGGCTTCAACAGCTTACAAACATGAAATTCCGAAGTGTGGA	1160
Db	1230	GACCCACACGTC-----GTTGACCTCTGACTGACCTAACAA	1271
QY	1161	GCAGTCTTCAGACGAGACACACAGCTTACATTTTGAGAAAGTGCACCAATCTCCAGGGGGCCA	1220
Db	1272	ACAGAGCATCTCATTTCCCGAGGGATTGAGGACTCTGTCCTCCCATCTTTCAGGAAATGC	1331
QY	1221	GAGCTGCCGCGGAGGAAAGAAACCGAAGTTTTCGCGCGGGAGCGCGGGTGCACCCGAG	1280
Db	1332	CTTTTACTCGCTCTCTCCCTCCACTCTGTGTCCTCCCTGACTGCACAGATTTAAGACAG	1391
QY	1281	CCCTGCCCGGAGGAAAGCCCCCGGGTAAACAAGCGCGGACGAGCGCGCTGGGGCC	1340
Db	1392	CAGCAGCTCCCTCACAGATCTCCAGGTTTCTTAAGTCAGAGTGCAGCTGACAGAGGCG	1451
QY	1341	CACGAGGAAAGAGAGAG--TGGTTAAGATAGACCCAGCAGAGATTAACCTCAGCC	1397
Db	1452	CAGTGGAGCAATGGCTGGAGGCCCCGACACATCGGCCAGGCCCTCGGGCCCCCGC	1511
QY	1398	CCGCGACGCAAGACAGAGCTCCCTGACTGGCAGTCTTCACCTTTCGAGCTGTGTGAA	1457
Db	1512	ACTCAGGCTGTGTCCAGGCTTTTCCAGAAAGTCTCAGGCTTCTCCCTCAGCCGGA	1571
QY	1458	CTCGCAGGGGGGAATGGCCCTCTGCACAGTACCTTGTCTCCCACTGAGCCCCGACAC	1517
Db	1572	GAGCCCCCTGATGTGAGCCCTGCAATTCAGATGCTT-----TGGGCCAGGTGA	1622
QY	1518	CAAGCAGGCTCTCCCTTGATCAGCGCGCTGATAGACAGGCTGCCCCGGAGGGA	1577
Db	1623	ACAGGGCTCATCTCCCATCTCTGAGCCCTGTCTCTCAGTATGCTGGCGGAGCGGGAATGA	1682
QY	1578	CGATGAGGATGAGGGCCGCGAGGAAGAGATTCACCACTGATTAAGCGTATCTCATAGCTAA	1637
Db	1683	CTGGAG--GAGCCAGACACAAAGCGGTGCTGACAGACGAGGCTTACATTAATGCA	1739
QY	1638	GGAAGTGTCTACCAACGAGGCAATATCTGAAGATTCGAAGTATCACTTGTGTGTT	1697
Db	1740	AGAAATTCCTGCTACAGAACGAAACATCTCTMAAGATTATGAAATTTACCGTGTGTT	1799
QY	1698	TCAGAGCAGATGAGCAAAAGGAGCGCACATGCCGGAAGCATGAAGTCTCATATTC	1757
Db	1800	CCGCAAGCGAGTGTGAAGAGGAGCGCATGCTGCGCATCTGTATGACGTGCTCTTCTC	1859
QY	1758	GAATTTTGAACCTTTGCAAAATTTTCATCTAATTTTCTCAAGAAATTGAGCACACT	1817
Db	1860	CAACATGATCCCATCTATATGATTTCCACAGAGGCTTCTGTGGCAGAGGTGAGCAGAGCT	1919
QY	1818	TGCCCTGTGGAGAGCGCGCTCAATATGCCCAATCAGAG--ATTACCAAGAAATCGGCGA	1874
Db	1920	GGCACTCTGGGAAAGGGCCCTCCAAAGCCCAACAAGGAGCAGTATCAAGAAATCGGGGA	1979
QY	1875	TGTCATCTGAAGAACTTCAGGGCATGAAGCACCTGGCGGCTACCTGTGGAAGCACAG	1934
Db	1980	CATCTGTCTCAGAAACATGCGCCAGTTTAAAGAAATTTTACAGTACTTCCAAAGACATGA	2039
QY	1935	CGAGGCTCTGGAGGCCCTGAGGAATGGAATCAAGAGCTCCGCGGGCTGAGAACTTCTG	1994
Db	2040	CGAGGTCTTAACAGAACTGGAAGAGGCTACCAACGCTGTAAAGAAATTGAGAGCAGTGT	2099
QY	1995	CAGAGACTTTGAGCTGCAGAAAGTGTGTACTTACGCGTCAACACTTCTCTGCGGCC	2054
Db	2100	CAAGAGATTGTAGCTGCAGAAAGGTCTGTACTTCTCTCAACAGTTTCTGTGAAGCC	2159
QY	2055	ACTGCACCGGCTCATGCACTCAAGCAGAGTCTTGGAAGCGGTGTGCAACAACACCCGCC	2114
Db	2160	CATCACAAGGGTGTGTGCACTACCGCTGTGCTGCGCGCTTAAGCGGACATTAACGCC	2219
QY	2115	GAGCCACGCGCACTTCAGAGGACTGCGAGCGCTTTGGACAGATATCACGAGATGTGTGC	2174
Db	2220	CGGGCACCAATACATACGCTGATCTGCTCATGAGCGCCTGAAGGCCATCACAGAGGTGACAC	2279
QY	2175	ACAGCTCCACGGTACGATGATCAAGATGAGAAATTTCCAGAAAGCTGCACGAATCAAGAA	2234

Db	2280	CACACTACAGCACATTCTCATCCGGCTGGAGAACTTGCAGAAAGCTACAGAGCTGCGACG	2339
Oy	2235	AGATTTGATTTGGCATTTGACAAATCTTGTGGTTCCGGAAAGGAGTTTCATCCGTCTGGGCA	2294
Db	2240	GGACCTGTGGGATATGAGAACCTCATTTGCTCTGGCAGGGAGTTTCATCCGTGAGGGCTG	2399
Oy	2295	CTTACAGCAAGCTCTCGGGAGAGGGGCTCCAGACAGGCACTGTTCTTCTGTTCAACGACGT	2354
Db	2400	CCTTCCACAGCTCACCMAAAGGGCTTGCAGCAGAGAGATGTTTTTCTGTCTCAGATAT	2459
Oy	2355	CTGTCTATTACAGAGCGGGGGCTGACGGCTCTCCATCAGTTTAAAGTCCAGGGCAGCT	2414
Db	2460	GTTGTCTGTATACAGAGAAAGAGTTGCAAGGACAGCACTTCCGAAATCCGGGGCTTCTCT	2519
Oy	2415	CCCGCTCTATGCGCATGACGATTGAGGAGAGCAAGACGAGTGGGGGTTGCCCTCATGCTCT	2474
Db	2520	TCCCTCTCAAGGCACTCTGTTGGAGAAAGTATATACAGTGTCTGTCTTCCACACTGTTT	2579
Oy	2475	GACCTCTCGGGGCGCAGCGGAGTCCATCATCTGTGCGCGCACTTCTCGGTCCGATGGA	2534
Db	2580	CACCATCTACGCGGCTCAGAAAAACAATCTGTGTGGCACCGACACTCCGGCTGGAGAAAG	2639
Oy	2535	GAAGTGGGTTGAGGAGATCCAGATGGCCATTGACCTGGCGGAGAGAGACACACGCCCGC	2594
Db	2640	GAAGTGGATGTGGAGCTTAACCTCCGCTGATCCAGACACCAAGATGGCGGTGACACGGC	2699
Oy	2595	CCCTGAGTTCTTGGCCAGCA-----GCCCCCTGTACAAACAGTCCCTGTATGGAAGCAC	2648
Db	2700	CCTGTCACTCCAGGCGCCGACCTGTGTGCACTGTCTCCCCAGATCCCCACAGAGATATC	2759
Oy	2649	CGCGGCTGACAGAGTCAAGAGATGACCTGAGCGCTCGCGCACATGCTGAGCGCCA	2708
Db	2760	TCTG-----GAGCAGAGATCAGAAAGATGATGCTCGGGGTTTCGCGAGCTCCTGGAGGGCA	2816
Oy	2709	GGCCCGGCAACCGGGGCAACAAATGATGCACTGTGTGTGGCAACCGCAACACACGCTTC	2768
Db	2817	TGGCCAGCACCGGGGCCMAACACCAATGACGTTGTGTGTGATACCGGAACACACGCTGTC	2876
Oy	2769	CATGTGTGACTTTCAGCATCGCAGTGGAGAAATCAGTTGTCTGGAAAACTGTGAGAAATT	2828
Db	2877	CAGGGCAGACACAGTGTGACGCTGTGAGAACCAAGCTTTCAAGATATCTGCTTAAGAAATT	2936
Oy	2829	CAAAAACAGCAACGGGTGGCAGAACTGTGGTGTGTTCAAAACTTCTGCTGTTCTT	2888
Db	2937	CAAAAACAGTCAATGGCTGGCAGAAAGTGTGGGTGCTTTTACAACTTCTGTGTTGTTCTT	2996
Oy	2889	CTTCAAAATCACACAGGACATATCATCTCCCTTGGCCAGCTGCTCTGCTCGGCTTACTGCT	2948
Db	2997	CTTCAAAATCTATCAGATGATCTAACCTCAGCTCCAGCTCCCGGCTGCGGCTTCAAGGT	3056
Oy	2949	CACATATCCCTCTGAGTGTCCGAGAAATCATCCAGAAACATACGTTTCAAGCTGCACCTTCAA	3008
Db	3057	GACATATCCCGAGGAGGCGGATGGCATATACAAACATATGTTTTCAGCTTCCAGTTTCA	3116
Oy	3009	GTCCTCAGTCTTACTTCTCAGGGCGGAAAGCAGATACAGTTTCCAAAGGTGATGAGAT	3068
Db	3117	ATCCCAAGTCTACTTCTTCCGGGCTGAGAGCAAGTACACATTTGAAAGGTGGAATGAGGT	3176
Oy	3069	GATCCGAGTGCACCAAGCTTGTGCTTCGGGACCCC	3103
Db	3177	GATCCAGGGGGCCAGCAGCTCAGCCGGGAGGGCCC	3211
RESULT 10			
ACD92278/c			
ID ACD92278 standard; cDNA; 582 BP.			
ACD92278;			
AC			
AC			
DT 23-SEP-2003 (first entry)			
XX			
XX Human colon cancer cell expressed cDNA #690.			

XX Open reading frame detection; genome sequencing; colon cancer;
 KW breast cancer; population genome analysis; genetic shift; cancer;
 KW antibiotic resistance; antibiotic non-tolerance; congenital disease;
 KW agriculture; food crop genome; resistance gene; retrovirus;
 KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
 KW gene; ss.
 XX
 OS Homo sapiens.
 PN US2002155438-A1.
 XX
 PD 24-OCT-2002.
 XX
 PF 27-SEP-1999; 99US-00406117.
 XX
 PR 20-NOV-1998; 98US-00196716.
 XX
 PA (SIMP/) SIMPSON A J G.
 PA (NETO/) NETO E D.
 PA (BREN/) BRENTANI R R.
 XX
 PI Simpson AJG, Neto ED, Brentani RR;
 DR WPI; 2003-182626/18.
 XX
 DR Determining open reading frames of genome of an organism e.g. a human
 PT suffering from cancer involves use of single oligonucleotide primer at
 PT low stringency for preparing single-stranded cDNA from mRNA of
 PT individual.
 XX
 PS Example 9; Page 105; 959pp; English.
 XX
 CC The invention describes a method of determining open reading frames in
 CC the genome of organism, comprising contacting mRNA from cell of organism
 CC with a single oligonucleotide primer (1) at low stringency, preparing
 CC single-stranded cDNA by reverse transcribing mRNA with (1), amplifying
 CC cDNA, sequencing the product, and repeating the contacting, preparing
 CC and amplifying steps with different primers and sequencing resulting
 CC nucleic acids. The method is useful for: determining that a known
 CC nucleotide sequence from a genome of an organism corresponds to a
 CC nucleotide sequence of an open reading frame; for preparing a contig,
 CC nucleic acid molecule from a genome of an organism; and for sequencing
 CC all or part of a genome of an organism. mRNA is obtained from mammalian
 CC or human cell which is associated with a pathological condition e.g. a
 CC colon cancer or breast cancer cell. The method is useful for analyses of
 CC populations of subjects and can be used to carry out genetic analyses of
 CC large or small populations. Further, it can be used to study living
 CC systems to determine if, e.g. there have been genetic shifts which render
 CC an individual or population more or less likely to be afflicted with
 CC diseases such as cancer, to determine antibiotic resistance or non-
 CC tolerance, and so forth. The method can also be used in the study of
 CC congenital diseases, and the risk of affliction to a fetus, as well as
 CC the study of whether the conditions are likely to be passed to offspring
 CC through ova or sperm. The analyses for pathological conditions can be
 CC carried out in all animals, plants, birds, fish, etc. Using this method,
 CC in the area of agriculture, for example the genomes of food crops can be
 CC studied to determine if resistance genes are present, defects in plant
 CC genomes can also be studied in this way. Similarly, the method permits
 CC determination of the pathogens which integrate into the genome, such as
 CC retroviruses and other integrating viruses such as influenza virus, have
 CC undergone shifts or mutations, which may require different approaches to
 CC therapy. This method is also applied to eukaryotic pathogens, such as
 CC trypanosomes, different types of Plasmodium, etc. The method essentially
 CC eliminates sequencing of non-coding portions. This sequence represents a
 CC polynucleotide isolated from human colon cancer cell cDNA library
 CC
 XX
 SQ Sequence 582 BP; 109 A; 147 C; 185 G; 140 T; 0 U; 1 Other;

Query Match 17.6%; Score 550.4; DB 10; Length 582;
 Best Local Similarity 99.6%; Pred. No. 1.3e-136;
 Matches 551; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2583 CAGCAGCCCCCGCCCTGAGTTCCTGCGCAGCAGCCCCCTGACAAAGTCCCTGATGA 2642
 DB |||||
 QY 571 CAGCAGCCCCCGCCCTTGAAGTTCCTGCGCAGCAGCCCCCTGACAAAGTCCCTGATGA 512
 DB |||||
 QY 2643 AGCCACCGCGGCTGACCGAGATCAGAGATGACTGAGCCGCTTGCGCAATCGCTGGA 2702
 DB |||||
 QY 511 AGCCACCGCGGCTGACCGAGATCAGAGATGACTGAGCCGCTTGCGCAATCGCTGGA 452
 DB |||||
 QY 2703 GCGCAGGCGCCCGCAGCGCGGCAACAGATGGTGCAGCTGTGTGAGCCGCAACCCAG 2762
 DB |||||
 QY 451 GCGCAGGCGCCCGCAGCGCGGCAACAGATGGTGCAGCTGTGTGAGCCGCAACCCAG 392
 DB |||||
 QY 2763 CGTCTCAGTGGTGAATTAGCATGCACTGAGAGATGATGTTGTGAGAACTGTGAG 2822
 DB |||||
 QY 391 CGTCTCAGTGGTGAATTAGCATGCACTGAGAGATGATGTTGTGAGAACTGTGAG 332
 DB |||||
 QY 2823 GAAATTCAAAAACAGCAACGGGTGGCAGAGAGTGTGGTGTTCACAAAATTGTGCT 2882
 DB |||||
 QY 331 GAAATTCAAAAACAGCAACGGGTGGCAGAGAGTGTGGTGTTCACAAAATTGTGCT 272
 DB |||||
 QY 2883 GTTCTTCTACAAATACACACAGAGATCATCCCTGCGCAGCTGCTGTGCGCTA 2942
 DB |||||
 QY 271 GTTCTTCTACAAATACACACAGAGATCATCCCTGCGCAGCTGCTGTGCGCTA 212
 DB |||||
 QY 2943 CTGCTCAACATCCCTCTTGAAGTCCGAGAACATCCAGAAAGCTACGTTCAGCTGCA 3002
 DB |||||
 QY 211 CTGCTCAACATCCCTCTTGAAGTCCGAGAACATCCAGAAAGCTACGTTCAGCTGCA 152
 DB |||||
 QY 3003 CTTCAAGTCCCAACGCTACTACTTTCAGAGCGGAAAGCGAGTACAGTTTGAAGGTGAT 3062
 DB |||||
 QY 151 CTTCAAGTCCCAACGCTACTACTTTCAGAGCGGAAAGCGAGTACAGTTTGAAGGTGAT 92
 DB |||||
 QY 3063 GGAAGTGAATCCGACATGCGACACAGGCTGCTGCGGACCCGACGTTGAGCCACAAAGA 3122
 DB |||||
 QY 91 GGAAGTGAATCCGACATGCGACACAGGCTGCTGCGGACCCGACGTTGAGCCACAAAGA 32
 DB |||||
 QY 3123 GTCTCTGTGTAT 3135
 DB |||||
 QY 31 GTCTCTGTGTAT 19
 DB |||||

RESULT 11
 ACD94832/c
 ID ACD94832 standard; cDNA; 582 BP.
 XX
 AC ACD94832;
 XX
 DT 23-SEP-2003 (first entry)
 XX
 DE Human colon cancer cell expressed cDNA #3244.
 XX
 KW Open reading frame detection; genome sequencing; colon cancer;
 KW breast cancer; population genome analysis; genetic shift; cancer;
 KW antibiotic resistance; antibiotic non-tolerance; congenital disease;
 KW agriculture; food crop genome; resistance gene; retrovirus;
 KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2002155438-A1.
 XX
 PD 24-OCT-2002.
 XX
 PF 27-SEP-1999; 99US-00406117.
 XX
 PR 20-NOV-1998; 98US-00196716.
 XX
 PA (SIMP/) SIMPSON A J G.
 PA (NETO/) NETO E D.
 PA (BREN/) BRENTANI R R.
 XX
 PI Simpson AJG, Neto ED, Brentani RR;

XX WPI; 2003-182626/18.

DR Determining open reading frames of genome of an organism e.g. a human
XX suffering from cancer involves use of single oligonucleotide primer at
PT low stringency for preparing single-stranded cDNA from mRNA of
PT individual.

XX Example 9; Page 477; 959pp; English.

XX The invention describes a method of determining open reading frames in
CC the genome of organism, comprising contacting mRNA from cell of organism
CC with a single oligonucleotide primer (I) at low stringency, preparing
CC single-stranded cDNA by reverse transcribing mRNA with (II), amplifying
CC cDNA, sequencing the product, and repeating the contacting, preparing
CC and amplifying steps with different primers and sequencing, resulting
CC nucleic acids. The method is useful for: determining that a known
CC nucleotide sequence from a genome of an organism corresponds to a
CC nucleic acid molecule of an open reading frame; for preparing a contig.
CC all or part of a genome of an organism. mRNA is obtained from mammalian
CC or human cell which is associated with a pathological condition e.g. a
CC colon cancer or breast cancer cell. The method is useful for analyses of
CC populations of subjects and can be used to carry out genetic analyses of
CC large or small populations. Further, it can be used to study living
CC systems to determine if, e.g. there have been genetic shifts which render
CC an individual or population more or less likely to be afflicted with
CC diseases such as cancer, to determine antibiotic resistance or non-
CC tolerance, and so forth. The method can also be used in the study of
CC congenital diseases, and the risk of affliction to a fetus, as well as
CC the study of whether the conditions are likely to be passed to offspring
CC through ova or sperm. The analyses for pathological conditions can be
CC carried out in all animals, plants, birds, fish, etc. Using this method,
CC in the area of agriculture, for example the genomes of food crops can be
CC studied to determine if resistance genes are present, defects in plant
CC genomes can also be studied in this way. Similarly, the method permits
CC determination of the pathogens which integrate into the genome, such as
CC retroviruses and other integrating viruses such as influenza virus, have
CC undergone shifts or mutations, which may require different approaches to
CC therapy. This method is also applied to eukaryotic pathogens, such as
CC trypanosomes, different types of Plasmodium, etc. The method essentially
CC eliminates sequencing of non-coding portions. This sequence represents a
CC polynucleotide isolated from human colon cancer cell cDNA library
XX

SO Sequence 582 BP; 109 A; 147 C; 185 G; 140 T; 0 U; 1 Other;

Query Match 17.6%; Score 550.4; DB 10; Length 582;

Best Local Similarity 99.6%; Pred. No. 1.3e-136;

Matches 551; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2583 CAGCAGCCCCCGCTGAGTTCCTGGCCAGCAGCCCCCGCTGACACAGTCCCTGATGA 2642
DB 571 CAGCAGCCCCCGCTGAGTTCCTGGCCAGCAGCCCCCGCTGACACAGTCCCTGATGA 512
QY 2643 AGCCACCGCGGCTGACAGAGTCAAGATGACAGGCGCTGCGGACATCGGTGA 2702
DB 511 AGCCACCGCGGCTGACAGAGTCAAGATGACAGGCGCTGCGGACATCGGTGA 452
QY 2703 GCGCCAGGCCCGCAGCAGCGGCAACAATGTGTGCTGTGACCGCAACACAG 2762
DB 451 GCGCCAGGCCCGCAGCAGCGGCAACAATGTGTGCTGTGACCGCAACACAG 392
QY 2763 CGTTCATGATGATTCAGCATTCGAGTGAAGATCAGTTGTCTGGAACCTGCTGAG 2822
DB 391 CGTTCATGATGATTCAGCATTCGAGTGAAGATCAGTTGTCTGGAACCTGCTGAG 332
QY 2823 GAAATTCAAAACAGCAACGCGGTGCGAGAAGCTGTGGGTGTTCAAACTTCGCT 2882
DB 331 GAAATTCAAAACAGCAACGCGGTGCGAGAAGCTGTGGGTGTTCAAACTTCGCT 272
QY 2883 GTTCTTCTAACAATCAACACAGCAACATATCCCTTGGCCAGGCTGCTGTGCGCTA 2942
DB 271 GTTCTTCTAACAATCAACACAGCAACATATCCCTTGGCCAGGCTGCTGTGCGCTA 212

QY 2943 CTGCTCACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTAGCTGTCAAGTCA 3002
DB 211 CTGCTCACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTAGCTGTCAAGTCA 152
QY 3003 CTTCAGTCCCAAGTCTACTACTTTCAGGCGGAAACGAGTACACGTTGGAAGTGAT 3062
DB 151 CTTCAGTCCCAAGTCTACTACTTTCAGGCGGAAACGAGTACACGTTGGAAGTGAT 92
QY 3063 GGAAGTATCCGAGTCCGAGCAGCAGCTGTGCGACCCACAGTTGAGCCACAAGA 3122
DB 91 GGAAGTATCCGAGTCCGAGCAGCAGCTGTGCGACCCACAGTTGAGCCACAAGA 32
QY 3123 GTTCTTGTGTAT 3135
DB 31 GTTCTTGTGTAT 19

RESULT 12

ID AD063863 standard; cDNA; 3200 BP.

XX AD063863;

DT 07-OCT-2004 (first entry)

DE Novel human cDNA sequence #1024.

XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KM cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KM neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
XX cancer.

OS Homo sapiens.

PN EP1440981-A2.

PD 28-UTL-2004.

XX 21-JAN-2004; 2004EP-00001196.

XX 21-JAN-2003; 2003JP-00102206.

PR 09-MAY-2003; 2003JP-00113392.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;

DR WPI; 2004-535376/52.
P-PSDB; AD066051.

PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PI Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1, SEQ ID NO 1024; 2449pp; English.

XX The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.

SO Sequence 3200 BP; 951 A; 659 C; 718 G; 872 T; 0 U; 0 Other;

Query Match 14.3%; Score 447; DB 12; Length 3200;

Best Local Similarity 65.5%; Pred. No. 1.6e-108;

Matches 654; Conservative 0; Mismatches 345; Indels 0; Gaps 0;

QY CCCAACACCTTGAGGAAAACTCGTGTCAATCAAAATCCAGATGCTGATGACACCCAGGA 155
DB CCCAACGGCTCGCTGATGATGCTATTAATAAGGAGAGTTTGGATGATCCAGAA 219
QY GGCATTGTAAGTTCCACAAAGAGCTCTGGAAGGTGCTGCTGATGACGTTGCAACA 215
DB GATTTTGTGGTTGATCAAAAGATCATCCGGAAAGGATGTTTAACTGATGATGCGCA 279
QY CCGCAACCTCGTGAAGGTGACTATTTGGCCCTGAGTTTCCGATCACAAAAGATCAC 275
DB TCTAAATCTTCTGAAAAGAAATATTTGGATTGAAATCTGACGACCTTCGAAATATA 339
QY GGTGTGCTGATATCTCTAAAACCATTTGAAACAGATTGAAGGCCAAAGACGTTGT 335
DB TGTGTGCTGAGCTTTTGAAGCCCATTAACAAAGAGGTAATAAAATCTTAAGAGATTGT 399
QY TGTTAAGTTTGTGTGAAATTTCTTTCCGCTGACACACACAACTCCAAAGAAATCTAC 395
DB TTTCAATTTATGTGTAAATTTTCCAGTGAACCTGAGCAATCTGCGGAAAGAACTTAC 459
QY AAGGTACCTGTCGAGCTGAGGTGAAGAGCACTTGGCTCAAGGAGGTTGACGTTAA 455
DB AAGGTATCTTTTACTCTTCAATTAAGAAAGATTTGGCTTAGGAAGGCTTCATGACG 519
QY TGAACACAGCGAGCTCTCTTGAATTCACACATTTGCAATCTGATGAGATTGGGATTTGA 515
DB TGAACATGTAAGCGCTTATGATGATCTCAATCTTACATGAGAACTTGAAGACTTTCA 579
QY TGAAGCTTGGACAGAGACCTTAGCAAAAATAATACATCTTACAGCAAGACGCACT 575
DB TGAAGAAACAGATGAGAACATCTGGACAAACTCGGTCTTACCAAAACCAAGACTGTT 639
QY AGAGGACAAATTCGTGAATTTCAACATTAACCATTTGAGCAAAACCCAGAAATCAGA 635
DB AGAGGGACAGATCATGCACTTTCATAGAAAGCAATTTGGAGAGAGCCGAGTAATCTGA 699
QY TTTCCAGCTCTAGAGATTCCTCGGCTAGAGATGATGAAATCCGTTACACCCGCGC 695
DB CATTCGTACTAGCATAGCAAGAAAGCTGGATATGATGGATCAGGCTCACCCTG 759
QY CAAAGACAGGAAAGGACAGAAATCAATTTGGCCCTTGGCAACAGGAAATCTAGTGT 755
DB CAGTATGTGTGAAGGATGCAATTCACCTGGCTGTTGTCACATGAGGATCTGTGT 819
QY TCAGGGTTTCATTAAGATCAATGCTTCAACTGAGGCGCAAGTGGGAGCTGAGCTTCAA 815
DB ACGGGAAATACAAAGATCAATATCTTTAACTGGGCTAAATCCGAAAGTTAATTTTAA 879
QY GAGAAAGCGCTTTCATCAAGCTCCGACAGATGCAATGATGCTGCTACAGGATACCTT 875
DB GAGAAAGCATTTTTCATCAAACTTCAATGCTCAATATCTTGTGTGTGCAAGGATACCTT 939
QY GGAATTTCTGATGCGCAGTGGGATTTCTGCAAGTCTTCTGGAATATCTGTGTGAACA 935
DB GGAGTTCAACATGCGCAGGAGTCCGCAAGGCTTCTGGAAGACTGTGTGGAATA 999
QY TCATGCTCTTCTTAAAGCTTTTGAAGAGCCCAAAACCAAGCCCAAGCCGCTCTTTAG 995
DB CATGCTTTCTTCAGGCTTTCGGAAGAGCCCAATTAAGCCCAAAACCTTACTCTGCGAG 1059
QY CCGGGGGTCATCATTTCCGTTCAAGTGTGGACTCGAAGACAGGTTCTGACTATGTTAA 1055
DB CAAGGGTTCAAGTTTCGCTATATAGTGAAGAACCCAAAGCAACTTTGGAATATGAGAG 1119
QY AGAAGAGAGACATTAAGAGGTGCAATTTGAAGAAAGCA 1094
DB AAAAGGAGGCTGAAAGAGCTTGCATTTGAAGAAAGCA 1158

RESULT 13
ACL61462
ID ACL61462 standard; cDNA, 435 BP.
XX

AC ACL61462;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human colon cancer differentially expressed polynucleotide, SEQ ID:7597.
XX
KW Differential expression; diagnosis; therapy; drug screening; cancer;
KM neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
XX
OS Homo sapiens.
XX
PN WO2005000087-A2.
XX
PD 06-JAN-2005.
XX
PF 13-MAY-2004; 2004WO-US015421.
XX
PR 03-JUN-2003; 2003US-0475872P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Randazzo F, Moler E, Escobedo J, Garcia PD;
XX
DR WPI; 2005-075421/08.
XX
PT New isolated polynucleotides, which are differentially expressed in colon
PT cancer cell, useful for treating cancer, e.g. colon cancer, breast
PT cancer, or pancreatic cancer.
XX
XX
XX Claim 1; SEQ ID NO 7597; 97bp; English.
XX
CC The invention relates to 9672 polynucleotides (ACLS3866-ACL61537) which
CC are differentially expressed in colon cancer cells. The invention also
CC relates to vectors and host cells comprising a differentially expressed
CC polynucleotide of the invention; a method for detecting a cancerous cell
CC by detection of a gene product of the polynucleotides; a method for
CC inhibiting a cancerous phenotype of a cell by inhibiting a gene product
CC of the polynucleotides; a method of treating an individual with cancer by
CC administration of a modulator of a gene product of the polynucleotides;
CC and an isolated antibody that specifically binds to a polypeptide encoded
CC by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
CC antibodies, and methods are useful for the detection of cancerous cells;
CC for the diagnosis, prognosis and management of cancer; for the
CC identification of agents that modulate the phenotype of cancerous cells;
CC and for the treatment of cancer, especially colon cancer and metastasized
CC colon cancer, but also breast or pancreatic cancer. The polynucleotides
CC are also useful as a source of probes or primers for use in diagnostic
CC methods. The differentially expressed polynucleotides or their encoded
CC proteins can additionally be used as vaccines to modulate primary immune
CC responses for the prevention or treatment of cancer. The present sequence
CC represents a specifically claimed polynucleotide which is differentially
CC expressed in colon cancer. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 435 BP; 100 A; 142 C; 115 G; 78 T; 0 U; 0 Other;

Query Match 13.4%; Score 420.8; DB 14; Length 435;
Best Local Similarity 99.3%; Pred. No. 6.1e-102;
Matches 433; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2540 GGGTTGAGAGACATCCAGATGCGCATTTGACCTGGCGGAGAAAGACGAGCCCGGCTGTG 2599
DB 1 GGGTTGAGAGACATCCAGA-GGCGATTTGACCTGGCGGAGAAAGACGAGCCCGGCTGTG 59
QY AGTTCTGAGCAGAGCCCGGCTGACAAAGTCCCGTATGAAAGCCACCGGCGCTGACC 2659
DB AGTTCTGAGCAGAGCCCGGCTGACAAAGTCCCGTATGAAAGCCACCGGCGCTGACC 119
QY 2660 AGGAGTCAAGAGATGACCTGAGCGGCTTCGCGACATCGCTGAGAGCGCAGGCCCGGCAAC 2719

```
Db 120 AGAGTCATAGATGACCTGAGCGCCCTCGGACATGCTGGAGCCGACGCCACC 179
Qy 2720 GCGGCAACAAATGCTGACGTGTGCTGGACACCGCAACACCGCTTCCATGTGACT 2779
Db 180 GCGGCAACAAATGCTGACGTGTGCTGGACACCGCAACACCGCTTCCATGTGACT 229
Qy 2780 TCAGCATCGAGTGGAGAAATCAGTTGCTGAAACCTGCTGAGAAATTCAAAACGCA 2839
Db 240 TCAGCATCGAGTGGAGAAATCAGTTGCTGAAACCTGCTGAGAAATTCAAAACGCA 299
Qy 2840 ACGGTCGCAAGAGCTGTGGGTGTGTTCAAACTTCTGCTGTTCTTACAAATCAC 2899
Db 300 ACGGTCGCAAGAGCTGTGGGTGTGTTCAAACTTCTGCTGTTCTTACAAATCAC 359
Qy 2900 ACCAGACATCATCTCCCTTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2959
Db 360 ACCAGACATCATCTCCCTTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
Qy 2960 CTGAGTCCGAGAACAT 2975
Db 420 CTGAGTCCGAGAACAT 435

RESULT 14
ACLS5653
ID ACLS5653 standard; cDNA; 637 BP.
XX ACLS5653;
XX
XX 24-MAR-2005 (first entry)
XX
DE Human colon cancer differentially expressed polynucleotide, SEQ ID:1788.
XX
XX Differential expression; diagnosis; therapy; drug screening; cancer;
XX neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
XX ss.
XX Homo sapiens.
XX
XX WO200500087-A2.
XX
XX 06-JAN-2005.
XX
XX 13-MAY-2004; 2004MO-US015421.
XX
XX 03-JUN-2003; 2003US-0475872P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Randazzo F, Moler E, Becobedo J, Garcia PD;
XX
XX WPI; 2005-075421/08.
XX
XX New isolated polynucleotides, which are differentially expressed in colon
XX cancer cell, useful for treating cancer, e.g. colon cancer, breast
XX cancer, or pancreatic cancer.
XX
XX Claim 1; SEQ ID NO 1788; 97bp; English.
XX
XX The invention relates to 9672 polynucleotides (ACLS5866-ACLS6337) which
XX are differentially expressed in colon cancer cells. The invention also
XX relates to vectors and host cells comprising a differentially expressed
XX polynucleotide of the invention; a method for detecting a cancerous cell
XX by detecting a gene product of the polynucleotides; a method for
XX inhibiting a cancerous phenotype of a cell by inhibiting a gene product
XX of the polynucleotides; a method of treating an individual with cancer by
XX administration of a modulator of a gene product of the polynucleotides;
XX and an isolated antibody that specifically binds to a polypeptide encoded
XX by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
XX antibodies, and methods are useful for the detection of cancerous cells;
XX for the diagnosis, prognosis and management of cancer; for the
XX identification of agents that modulate the phenotype of cancerous cells;
```

```
CC for the identification of therapeutic targets for cancer chemotherapy;
CC and for the treatment of cancer, especially colon cancer and metastasized
CC colon cancer, but also breast or pancreatic cancer. The polynucleotides
CC are also useful as a source of probes or primers for use in diagnostic
CC methods. The differentially expressed polynucleotides or their encoded
CC proteins can additionally be used as vaccines to modulate primary immune
CC responses for the prevention or treatment of cancer. The present sequence
CC represents a specifically claimed polynucleotide which is differentially
CC expressed in colon cancer. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 637 BP; 144 A; 192 C; 153 G; 148 T; 0 U; 0 Other;
XX
XX Query Match 12.9%; Score 403.4; DB 14; Length 637;
XX Best Local Similarity 99.8%; Pred. No. 3.5e-97;
XX Matches 404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
Qy 2731 ATGGTGACGTGTGCTGGACACCGCAACACGAGGTCTCCATGTGGAATTACAGATGCA 2790
Db 11 ATGGTGACGTGTGCTGGACACCGCAACACGAGGTCTCCATGTGGAATTACAGATGCA 70
Qy 2791 GTGAGAAATCAGTTGCTGGAACCTGCTGAGGAAATTCAAAACAGCAACGGGTGGCAG 2850
Db 71 GTGAGAAATCAGTTGCTGGAACCTGCTGAGGAAATTCAAAACAGCAACGGGTGGCAG 130
Qy 2851 AAGCTGTGGGTGTGTTCAAACTTCTGCTGTTCTTCAAAATCACACGAGACAT 2910
Db 131 AAGCTGTGGGTGTGTTCAAACTTCTGCTGTTCTTCAAAATCACACGAGACAT 190
Qy 2911 CATCCCTTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2970
Db 191 CATCCCTTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 250
Qy 2971 AACATCCAGAAAGACTACGTTCAAGCTGCACTTAAGTCCACAGCTCTACTTACAG 3030
Db 251 AACATCCAGAAAGACTACGTTCAAGCTGCACTTAAGTCCACAGCTCTACTTACAG 310
Qy 3031 GCGGAAAGCGAGTACACGTTTCGAAGGTGATGAAAGTATCCGACGTCCACAGCTCT 3090
Db 311 GCGGAAAGCGAGTACACGTTTCGAAGGTGATGAAAGTATCCGACGTCCACAGCTCT 370
Qy 3091 GCCTCGGACCCCGACGTTTGAGCCCAAAAGATCTCTGTGAT 3135
Db 371 GCCTCGGACCCCGACGTTTGAGCCCAAAAGATCTCTGTGAT 415

RESULT 15
AAS27016
ID AAS27016 standard; cDNA; 1718 BP.
XX AAS27016;
XX
XX 07-NOV-2001 (first entry)
XX
XX cDNA encoding novel signal transduction pathway protein, Seq ID 51.
XX
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
XX anti-inflammatory; anti-HIV; antibacterial; anti-inflammatory; cancer;
XX immune system disorder; rheumatoid arthritis; inflammatory condition;
XX organ transplant rejection; infection; hepatitis C; blood disorder;
XX sickle cell anemia; hyperproliferative disorder; Gaucher's disease;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX chromosomal abnormality; Down syndrome; ischemia; renal disorder;
XX cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
XX acquired immune deficiency syndrome.
XX
XX Homo sapiens.
XX
XX OS
XX PN WO200154733-A1.
XX
```

PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001312.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-022547P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226779P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-023497P.
PR 25-SEP-2000; 2000US-023498P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-023637P.
PR 29-SEP-2000; 2000US-023637P.

PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.

(HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465460/50.

DR P-PSDB; AAU1099.

XX Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders.

PS Claim 1; SEQ ID NO 51; 880bp; English.

XX The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, hemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative disorders
 CC (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease), chromosomal abnormalities
 CC (Down syndrome), ischemic injury (e.g. stroke), renal disorders (e.g.
 CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),
 CC respiratory disorders, dermatological disorders, in wound healing,
 CC epithelial cell proliferation, endocrine disorders (e.g. Addison's
 CC disease), reproductive system disorders, gastrointestinal disorder
 CC (inflammatory disorders), liver disorders (cirrhosis), as stimulators of
 CC B-cell responsiveness to pathogens, activators of T-cells, to induce
 CC higher affinity antibodies, and as a means to induce tumour proliferation
 CC in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS2676-
 CC AAS27850 represent novel signal transduction pathway protein coding
 CC sequences and PCR primers of the invention

Query Match 12.8%; Score 402.4; DB 4; Length 1718;

Best Local Similarity 67.1%; Pred. No. 1, 1e-96;

Matches 603; Conservative 0; Mismatches 286; Indels 9; Gaps 2;

QY 2212 CAGAGCTGACGAGACTCAGAGAAATTTGATGGCATTGATGATCTGTGTTCCGGGA 2271
 Db 13 CAGAGCTGACGAGACTCAGAGAAATTTGATGGCATTGATGATCTGTGTTCCGGGA 72
 QY 2272 AGGAGTTTCTCGTGTGGGACCTTCAGAGCTCTCGGGGAAAGGGCTCCAGACGC 2331
 Db 73 AGGAGTTTCTCGTGTGGGACCTTCAGAGCTCTCGGGGAAAGGGCTCCAGACGC 132
 QY 2332 ATGTTCTTCTGTTCAACGACGTCCTATACAGAGCCGGGGCTGACGGCTCCAAT 2391
 Db 133 ATGTTCTTCTGTTCAACGACGTCCTATACAGAGCCGGGGCTGACGGCTCCAAT 192
 QY 2392 CAGTTTAAAGTCCACGGGAGCTCCGCTCTATGGCATGACGATTGAGAGAGAGAGAC 2451
 Db 193 CAGTTTAAAGTCCACGGGAGCTCCGCTCTATGGCATGACGATTGAGAGAGAGAGAC 252
 QY 2452 GAGTGGGGGGTCCCACTGACCTTCGAGGCGCAGGCGCATGTCATCATCTGAGCC 2511
 Db 253 GAGTGGGGGGTCCCACTGACCTTCGAGGCGCAGGCGCATGTCATCATCTGAGCC 312
 QY 2512 GCCAGTTTCTGATCCAGATGAGAAAGTGGTGAAGACATCCAGATGGCATTTGACCTG 2571
 Db 313 GCCAGTTTCTGATCCAGATGAGAAAGTGGTGAAGACATCCAGATGGCATTTGACCTG 372
 QY 2572 GCGGAG 2625
 Db 373 GCGGAG 432
 QY 2626 AACCAAGTCCCTGATGAGACCGCGCTGACCAAGAGTCAAGAGATGACCTGAGCGCC 2685
 Db 433 AACCAAGTCCCTGATGAGACCGCGCTGACCAAGAGTCAAGAGATGACCTGAGCGCC 489
 QY 2686 TCGCGACATCGGTGAG 2745
 Db 490 TCGCGACATCGGTGAG 549
 QY 2746 TGGCACCAGCAACCAAGCTCTCATGTTGAGATTCAGCATCGAGTGAAGATCAGTTG 2805

Db 550 TGGTACCGGAAACACAGCGGTGTCAGAGGACAGACAGTGCACGTGTCCAGAAACAGCTT 609
 QY 2806 TCTGAAACCTGCTGAGAGAAATTCAAAACAGCAAGGCTGGCAAGAGCTGTGGGTG 2865
 Db 610 TCAAGATATCTGCTAAGAAAGTTCAAAACAGTCATGCTGGCAAGAGCTGTGGGTG 669
 QY 2866 TTTCAAACTTCGCTGCTGCTTCTTCAAAATCAGCAGAGCAATATCCCTTGGCAGC 2925
 Db 670 TTTCAAACTTCGCTGCTGCTTCTTCAAAATCAGCAGAGCAATATCCCTTGGCAGC 729
 QY 2926 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2985
 Db 730 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789
 QY 2986 TACGTGTTCAAGTGAATTCAGATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3045
 Db 790 TATGTTTCAAGTGAATTCAGATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 849
 QY 3046 ACGTTGAAAGTGAATTCAGATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3103
 Db 850 ACGTTGAAAGTGAATTCAGATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 907

Search completed: May 22, 2006, 01:14:41
 Job time: 1233.93 sec

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2006, 00:47:45 ; Search time 13987.1 Seconds
(Without alignments)
12533.446 Million cell updates/sec

Title: US-09-555-342b-1_COPY_49_3183

Perfect score: 3135
Sequence: 1 atgggggaacacagacagag.....acaaagagctctctgtgtat 3135

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est3: *
3: gb_est4: *
4: gb_est5: *
5: gb_est6: *
6: gb_hic: *
7: gb_est2: *
8: gb_est7: *
9: gb_est8: *
10: gb_est9: *
11: gb_g881: *
12: gb_g882: *
13: gb_g883: *
14: gb_g884: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3133.4	99.9	3138	14	AY406167 Homo sapi
2	3055	97.4	4375	6	CR859100 Pongo pyg
3	2965.2	94.6	3138	14	AY406168 Pan trogl
4	2425	77.4	3147	14	AY406169 Mus muscu
5	1013.2	32.3	3743	6	AK171713 Mus muscu
6	1007.7	32.1	3874	6	AK050860 Mus muscu
7	957	30.5	968	4	BX364606 Mus muscu
8	940.2	30.0	1000	4	BX354921 BX354921
9	883.4	28.2	912	3	BQ881649 AGENCOURT
10	873.2	27.9	904	3	BQ931960 AGENCOURT
11	863.8	27.6	1026	2	BM551766 AGENCOURT
12	852.4	27.2	1072	4	BX354920 BX354920
13	842.4	26.9	3165	14	AY415515 Homo sapi
14	839.8	26.8	959	3	BUS39080 AGENCOURT
15	834	26.6	857	4	BX350128 BX350128
16	827	26.4	916	3	BUS84695 AGENCOURT
17	819.4	26.1	951	4	BX364605 BX364605
18	817.2	26.1	966	3	BUI46764 AGENCOURT
19	816.6	26.0	889	3	BUI74449 AGENCOURT

20	812.6	25.9	926	3	BQ706499	BQ706499 AGENCOURT
21	802.2	25.6	3168	14	AY415517	AY415517 Mus muscu
22	801.2	25.6	806	10	DR763338	DR763338 HESG4_151
23	797.8	25.4	899	4	BX327170	BX327170 BX327170
24	795.8	25.4	800	9	CX781945	CX781945 HESG3_10
25	789.4	25.2	1354	10	DV784833	DV784833 Hw liver
26	788.2	25.1	1043	3	BQ072025	BQ072025 AGENCOURT
27	784.8	25.0	1132	7	BE746268	BE746268 601579756
28	781.8	24.9	886	3	BQ672632	BQ672632 AGENCOURT
29	781	24.9	933	3	BQ950768	BQ950768 AGENCOURT
30	762	24.3	774	2	B1760610	B1760610 603044980
31	761.2	24.3	804	2	BG747792	BG747792 602705293
32	760.6	24.3	881	3	BUI63030	BUI63030 AGENCOURT
33	757.2	24.2	865	7	BE260677	BE260677 601154188
34	755.8	24.1	784	2	BG763918	BG763918 602736888
35	755.4	24.1	988	3	BQ674703	BQ674703 AGENCOURT
36	753	24.0	930	1	AL566821	AL566821 AL566821
37	751.8	24.0	952	3	BQ676994	BQ676994 AGENCOURT
38	748	23.9	944	3	BUI73111	BUI73111 AGENCOURT
39	745	23.8	861	2	BG764061	BG764061 602737071
40	730.4	23.3	745	8	CN281631	CN291631 170004710
41	726.2	23.2	901	2	BG475554	BG475554 602491544
42	725.6	23.1	3126	14	AY415516	AY415516 Pan trogl
43	719.2	22.9	942	3	BUI79876	BUI79876 AGENCOURT
44	716.4	22.9	741	8	CN291626	CN291626 170004245
45	716	22.8	736	2	BF793662	BF793662 602254351

ALIGNMENTS

RESULT 1
AY406167
LOCUS
DEFINITION
Homo sapiens FARP1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY406167.1 GI:39762141
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo
1 (bases 1 to 3138)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shteyn,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)

JOURNAL
PUBMED
14671302
2 (bases 1 to 3138)

AUTHORS
Clark,A.G., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shteyn,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

TITLE
JOURNAL
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
1..3138
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..3138
/gene="FARP1"
/locus_tag="HCM2467"

ORIGIN

Query Match 99.9%; Score 3133.4; DB 14; Length 3138;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGAAATATAGACGAGGCGGACCCCGATGACGACTGGGGGCCCCGGAATTTGG 60
DB 1 ATGGGAAATATAGACGAGGCGGACCCCGATGACGACTGGGGGCCCCGGAATTTGG 60
QY 61 GGGATAGTACCTTGGAGCGTGGACAGAAAGCGGCCCCCAACCTTCAGGAAAATCGTG 120
DB 61 GGGATAGTACCTTGGAGCGTGGACAGAAAGCGGCCCCCAACCTTCAGGAAAATCGTG 120
QY 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAGTTCCACAAGAGCT 180
DB 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAGTTCCACAAGAGCT 180
QY 181 CCGGGAGAGGTGCTGCTGATGATGACACCCAGAGAGCATTTGAGTTCCACAAGAGCT 240
DB 181 CCGGGAGAGGTGCTGCTGATGATGACACCCAGAGAGCATTTGAGTTCCACAAGAGCT 240
QY 241 TTTGGCTGAGTTTCTGATCAAAAAGATCAGCGGTGGCTGATCTCTAAAACC 300
DB 241 TTTGGCTGAGTTTCTGATCAAAAAGATCAGCGGTGGCTGATCTCTAAAACC 300
QY 301 ATTGTAAACAGATTAGAAAGGCCAAAGCAGTTGTTTAAAGTTTGTGTAATTTCTT 360
DB 301 ATTGTAAACAGATTAGAAAGGCCAAAGCAGTTGTTTAAAGTTTGTGTAATTTCTT 360
QY 361 CGGCTGACCAACAACAATCTCCAAAGAACTCAACAAGTACTGTTCCGCGCTGCAGGTG 420
DB 361 CGGCTGACCAACAACAATCTCCAAAGAACTCAACAAGTACTGTTCCGCGCTGCAGGTG 420
QY 421 AAGCAGACTTGGCTCAGAGCAGGTGACGTGTATGACACAGCGAGCTCTTGATT 480
DB 421 AAGCAGACTTGGCTCAGAGCAGGTGACGTGTATGACACAGCGAGCTCTTGATT 480
QY 481 TCACACATTGTGCAATCTGAGATTGGGGAATTTGATGAAAGCTTGACAGAGACCTTA 540
DB 481 TCACACATTGTGCAATCTGAGATTGGGGAATTTGATGAAAGCTTGACAGAGACCTTA 540
QY 541 GCAAAAATTAATATACATCTCTCAGCAAGACGCACTAGAGACAAATCGTGAATTTTAC 600
DB 541 GCAAAAATTAATATACATCTCTCAGCAAGACGCACTAGAGACAAATCGTGAATTTTAC 600
QY 601 CATTAACCACTTGGACAAACACAGCAGATCAGATTTCCAGCTCTAGAGATTGCCGT 660
DB 601 CATTAACCACTTGGACAAACACAGCAGATCAGATTTCCAGCTCTAGAGATTGCCGT 660
QY 661 CGGCTAGAGATGTATGGAATCCGATTGCAACCCGCGCAAGGACAGGAGCAAGAGATC 720
DB 661 CGGCTAGAGATGTATGGAATCCGATTGCAACCCGCGCAAGGAGCAAGAGATC 720
QY 721 AATCTGAGCGTTGCCAACACGGGAAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
DB 721 AATCTGAGCGTTGCCAACACGGGAAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
QY 781 TTCAACTGGGCGCAAGGTCGGAAGCTGAGCTTCAAGAGAAAGCGCTTTCTCAACAGCTC 840
DB 781 TTCAACTGGGCGCAAGGTCGGAAGCTGAGCTTCAAGAGAAAGCGCTTTCTCAACAGCTC 840
QY 841 CGGCGAGATGCAAAATAGTGGTAACAAGATACCTTGAATTTCTGATGGCAGTGGGAT 900
DB 841 CGGCGAGATGCAAAATAGTGGTAACAAGATACCTTGAATTTCTGATGGCAGTGGGAT 900
QY 901 TTTTGAAGTCTTTTGGAAAATCTGTGTGAACATATGCTTTTGAATTTTGA 960
DB 901 TTTTGAAGTCTTTTGGAAAATCTGTGTGAACATATGCTTTTGAATTTTGA 960
QY 961 GAGCCCAACCAAGCCCAAGCCGCTCTCTTAGCCGGGATCATCATTTGGATTGAGT 1020
DB 961 GAGCCCAACCAAGCCCAAGCCGCTCTCTTAGCCGGGATCATCATTTGGATTGAGT 1020
QY 1021 GGTGGAGCTCAGAAAGAGTTCTGACTATGTTAAAGAGAGACATAAGAGGTGACG 1080

DB 1021 GGTGGAGCTCAGAAACAGGTTCTGACTATGTTAAAGAGAGACATAAGAGGTGACG 1080
QY 1081 TTTGAAAGAGACACAGCAAGATTGATTTCTATCCGAGCCTTGCTCAACAGCTACAGAA 1140
DB 1081 TTTGAAAGAGACACAGCAAGATTGATTTCTATCCGAGCCTTGCTCAACAGCTACAGAA 1140
QY 1141 CTGAATTCGAAATGCTGAGAGCATCTCAGCAGAGCAGCCCTTACATTTGAGAGAGGT 1200
DB 1141 CTGAATTCGAAATGCTGAGAGCATCTCAGCAGAGCAGCCCTTACATTTGAGAGAGGT 1200
QY 1201 GCCGAATCTCCAGGGGGCCAGAGCTGCGCGCAGAGAAAGAACCGAAGTTTCCGCGGG 1260
DB 1201 GCCGAATCTCCAGGGGGCCAGAGCTGCGCGCAGAGAAAGAACCGAAGTTTCCGCGGG 1260
QY 1261 GAGCCGGGGTGGCACCCGAGCCCTGCGCGCAGAGAAAGCCCGGGGTAAACAAGCAGGG 1320
DB 1261 GAGCCGGGGTGGCACCCGAGCCCTGCGCGCAGAGAAAGCCCGGGGTAAACAAGCAGGG 1320
QY 1321 GACGAGCCGCTCGGCGCCACGAGAGAAAGAGAGGTGTTAAGATAGAACCCAG 1380
DB 1321 GACGAGCCGCTCGGCGCCACGAGAGAAAGAGAGGTGTTAAGATAGAACCCAG 1380
QY 1381 CAGAGTAACTTACAGCCCGCCAGCCCAAGCAGAGCTCTGACTGACAGTCTCACTT 1440
DB 1381 CAGAGTAACTTACAGCCCGCCAGCCCAAGCAGAGCTCTGACTGACAGTCTCACTT 1440
QY 1441 TCCGAGCTGCTGTGTAACCTCGAGGGGGAGTGGCCCTGCGCAAGTGAAGCTTGTGCC 1500
DB 1441 TCCGAGCTGCTGTGTAACCTCGAGGGGGAGTGGCCCTGCGCAAGTGAAGCTTGTGCC 1500
QY 1501 AACCTGAGCCCGGACACCAAGAGGCTCTCCCTTGATCAGCCGCTGCTGAATACAG 1560
DB 1501 AACCTGAGCCCGGACACCAAGAGGCTCTCCCTTGATCAGCCGCTGCTGAATACAG 1560
QY 1561 GCCTGCCCCCGGACGGAAGATGAGAGTGGGCGGAGAAAGATTTCCAACTGATAA 1620
DB 1561 GCCTGCCCCCGGACGGAAGATGAGAGTGGGCGGAGAAAGATTTCCAACTGATAA 1620
QY 1621 GGGTACTTATAGCTTAAGGAAGTGTATCACCGAGCCGAACATATCTGAAGATCTCGAA 1680
DB 1621 GGGTACTTATAGCTTAAGGAAGTGTATCACCGAGCCGAACATATCTGAAGATCTCGAA 1680
QY 1681 GTTATCACTTCGTGTTTCAAGACAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGT 1740
DB 1681 GTTATCACTTCGTGTTTCAAGACAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGTGAAGCAAGT 1740
QY 1741 AAAAGTCTATATTCGCCGAATTTTGAACCTTTGCAAAATTTCTAATTTTCTCAAG 1800
DB 1741 AAAAGTCTATATTCGCCGAATTTTGAACCTTTGCAAAATTTCTAATTTTCTCAAG 1800
QY 1801 GAAATTGAGCAACGATTCGCGCTGGGGAAGGCGCTCAAAATGCCCCAATTCAGAGTTAC 1860
DB 1801 GAAATTGAGCAACGATTCGCGCTGGGGAAGGCGCTCAAAATGCCCCAATTCAGAGTTAC 1860
QY 1861 CAAAGATCGGAGATGTATGCTGAAGAACATTCAGGGCATGAAGCACTGGGGCTCAC 1920
DB 1861 CAAAGATCGGAGATGTATGCTGAAGAACATTCAGGGCATGAAGCACTGGGGCTCAC 1920
QY 1921 CTGTGAGACACAGGAGGCTTTGAGAGGCTCTGAGAGATGGAATCAAGAGCTCCCGCGG 1980
DB 1921 CTGTGAGACACAGGAGGCTTTGAGAGGCTCTGAGAGATGGAATCAAGAGCTCCCGCGG 1980
QY 1981 CTGGAAGACTTTCGAGAGACTTTGAGCTGCAAGAGGTGTGTAACCTACCGCTCAACACC 2040
DB 1981 CTGGAAGACTTTCGAGAGACTTTGAGCTGCAAGAGGTGTGTAACCTACCGCTCAACACC 2040
QY 2041 TTTCTCTGCGGCACTGCAACCGGCTCATGACATCAAGAGAGTCTTGAAGCGGCTGTGC 2100
DB 2041 TTTCTCTGCGGCACTGCAACCGGCTCATGACATCAAGAGAGTCTTGAAGCGGCTGTGC 2100
QY 2101 AAAACACACCCGCGGAGCGGCTTCAAGGACTGCTGAGCGGCTTTGGCAGAGATC 2160

Db 2101 AAACACACCCCGGAGCCAGCCGCACTTCAGGAGCTGCGAGCGGCTTTGGCAGAGATC 2160
 Qy 2161 ACCGAGATGAGGACAGCTCCAGGTATCGATGATCAAGAGAAATTTCCAGAAAGTGG 2220
 Db 2161 ACCGAGATGAGGACAGCTCCAGGTATCGATGATCAAGAGAAATTTCCAGAAAGTGG 2220
 Qy 2221 CACGAACTCAGAAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2280
 Db 2221 CACGAACTCAGAAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2280
 Qy 2281 ATCCGCTTGGGACAGCTTCAGCAAGCTTCGCGGAAAGGGGCTCCAGCAGCGCATGTTCTTC 2340
 Db 2281 ATCCGCTTGGGACAGCTTCAGCAAGCTTCGCGGAAAGGGGCTCCAGCAGCGCATGTTCTTC 2340
 Qy 2341 CTGTTCAACAGCGCTCTGATATACAGACCGCGGGGCTGAGAGGGCTCCAAATCAATTTTAAA 2400
 Db 2341 CTGTTCAACAGCGCTCTGATATACAGACCGCGGGGCTGAGAGGGCTCCAAATCAATTTTAAA 2400
 Qy 2401 GTCCACGGGAGCTCCGCTCTATGAGATGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460
 Db 2401 GTCCACGGGAGCTCCGCTCTATGAGATGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460
 Qy 2461 GTGCCCCACCTGCTGACCTTCGCGGGGCAAGGAGATTCATCATGTTGGCCGCAATTTCT 2520
 Db 2461 GTGCCCCACCTGCTGACCTTCGCGGGGCAAGGAGATTCATCATGTTGGCCGCAATTTCT 2520
 Qy 2521 CGGTCGAGATGAGAGATGAGATGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2580
 Db 2521 CGGTCGAGATGAGAGATGAGATGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2580
 Qy 2581 AGCAGACGCCCGCCCTGAGATTCCTGGGCGAGAGCCCGCTGCAACAAGTCCCTGAT 2640
 Db 2581 AGCAGACGCCCGCCCTGAGATTCCTGGGCGAGAGCCCGCTGCAACAAGTCCCTGAT 2640
 Qy 2641 GAAAGCCACCGCGCTGACAGAGATCAGAGATGAGATTCAGATTCAGATTCAGATTCAGATTC 2700
 Db 2641 GAAAGCCACCGCGCTGACAGAGATCAGAGATGAGATTCAGATTCAGATTCAGATTCAGATTC 2700
 Qy 2701 GAGGCGCAGGCGCCGCGCAGCGGCGCAACAATGATGACGTTGCTGGCAGCCGCAACACC 2760
 Db 2701 GAGGCGCAGGCGCCGCGCAGCGGCGCAACAATGATGACGTTGCTGGCAGCCGCAACACC 2760
 Qy 2761 AGCGTCCATGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2820
 Db 2761 AGCGTCCATGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 2820
 Qy 2821 AGGAAATTCAAAAACAGCAACGAGTGGCAGAAAGCTGTGGTGTGTTCAAAAATTCTGC 2880
 Db 2821 AGGAAATTCAAAAACAGCAACGAGTGGCAGAAAGCTGTGGTGTGTTCAAAAATTCTGC 2880
 Qy 2881 CTGTTCTTCAAAATCAACACAGAGCAATATCCCTTGGCAGCTGCTCTGCTGGC 2940
 Db 2881 CTGTTCTTCAAAATCAACACAGAGCAATATCCCTTGGCAGCTGCTCTGCTGGC 2940
 Qy 2941 TACTCGCTCAGCATCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTTTCAAGCTG 3000
 Db 2941 TACTCGCTCAGCATCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTTTCAAGCTG 3000
 Qy 3001 CACTTCAGATTCAGCATCTTCAAGGCGGAAAGCGAGTACAGCTTGGAAAGGTGG 3060
 Db 3001 CACTTCAGATTCAGCATCTTCAAGGCGGAAAGCGAGTACAGCTTGGAAAGGTGG 3060
 Qy 3061 ATGGAAGTATCCGCAATGCAACAGCTTGGCTCGGAGCCCACTGTTGAGCCACAAA 3120
 Db 3061 ATGGAAGTATCCGCAATGCAACAGCTTGGCTCGGAGCCCACTGTTGAGCCACAAA 3120
 Qy 3121 GAGTCTCTGTGTAT 3135
 Db 3121 GAGTCTCTGTGTAT 3135

LOCUS CR859100 4775 bp mRNA linear HTC 12-NOV-2004
 DEFINITION Pongo pygmaeus mRNA; cDNA DKFZp469C092 (from clone DKFZp469C092).
 ACCESSION CR859100
 VERSION CR859100.1 GI:55729110
 KEYWORDS HTC.
 SOURCE Pongo pygmaeus (orangutan)
 ORGANISM Pongo pygmaeus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Pongo.
 REFERENCE 1 (bases 1 to 4775)
 AUTHORS Bioecker,H., Boecker,M., Brandt,P., Mewes,H.W., Weil,B., Amid,C., Oanger,A., Fobo,G., Han,M. and Wiemann,S.
 CONSRM The German cDNA Consortium
 TITLE Direct Submission
 JOURNAL Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
 COMMENT This clone (DKFZp469C092) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/ci.cgi?cloneID=DKFZp469C092
 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
 FEATURES
 source
 1..4775
 /organism="Pongo pygmaeus"
 /mol_type="mRNA"
 /db_xref="taxon:9600"
 /clone="DKFZp469C092"
 /issue_type="kidney"
 /clone_id="469 (synonym: pkid1). Vector pSport1_Sfi; host DH10B; sites SfiIa + SfiIb"
 /dev_stage="adult"
 /note="F8BM, RhogEF, and plectectrin domain protein 1 (Homo sapiens)"
 1..4775
 /gene="DKFZp469C092"
 100..3237
 /gene="DKFZp469C092"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAH91292.1"
 /db_xref="GI:55729111"
 /translation="MGSEIORPTPGSRILGAPENSGLSTLERGOKPPTPSGKLSIKI
 OMLDIOEAFEPVORAPGKTLDAVNCNHLNVGSGVLECPDHKKITWLLDLKPIV
 KOIRRKHVYVKKVFPFDPHTOLOBELTRYIFALQVOKDLAQGLTCDTDSALLI
 SHIVQSEIGDFDALDREHLAKNKYIPQDALBCKLVEFHNIHIGTPASDFOLBI
 ARRLBWYGLRLHAKOREGKILAVANTGILVFOGKTLNANMAAVRLSPRKRF
 LILKLPDANAVYDLEFLMASRDPCKSEFKIVLVEHAFRLFEERPKRPVLFSG
 SSFRSGRTQKQVLDVYKGGHKKVQPERKSHLSHLSASQPTLSVLSQOOS
 ASLTPREGAESPGSGSCCGCKEPKVPGRGSGSPSPSPSPSPAGNTOADAASAPES
 EEEVYVDRTQSKRPQRPQSTGSLTSGSPHLSLSVNSQGVANANTYLSNLPDTRQ
 ASPLISPLANDQACPTDDEBGRKRFPDKAVFIAKEVSTTERKYLKOLEVTSWF
 QSAVSEKDAPEALKSLIFENFEPRLKFNPLKEIQRALWEGSNADIRYORG
 DVMILKNIQGMKHLVHLMKSEBALLENIGKSRLLENCRDFEIOKVCYLPNPLFL
 LRLHLHMHYKQVLERLCKHPPSHADPDCRAALAGITEMVNLQGTMLKMNPOCL
 HELKDLIGIDNLVVRGRFIRLGSLSKLSGKLOQRMFLFNDVLLYSRGLTASND
 FYKHGOLPLYGMITKESBDEWGPCHCTILRQGSQITVAASSSEMEKVEDIQMAD
 LAEKNSSLAEFLASSPDPKSPDEATVADQESDDUSASRTSLERADPRGVTWAV
 CMHRNTSVSWPDSVAVENQLSGNLRLKPKNSGWLWVVFNFCLFFPKYSQDNDP
 LAGLPLGYSILTIPTSESENIHKDYVFKLHFKSHVYVFRASEYTFERMEVIRSAVSS
 ASRVHVSHSKESLVY"
 ORIGIN
 Query Match 97.4%; Score 3055; DB 6; Length 4775;
 Best Local Similarity 98.4%; Pred. No. 0;
 Matches 3085; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1 ATGGAGAAATAGAGAGAGGCGACCCAGATCAAGCTGGGGCCCCCGAAATTCG 60
Db 100 ATGGAGAAATAGAGAGAGGCGACCCAGATCAAGCTGGGGCCCCCGAAATTCG 159
QY 61 GGGATCAGTACCTTGGAACTGGACAGAGCCGCCCAACACTTCAGAGAAATCTCGTG 120
Db 160 GGGATCAGTACCTTGGAACTGGACAGAGCCGCCCAACACTTCAGAGAAATCTCGTG 219
QY 121 TCCATCAAAATCGAGATGCTGATGACACCCAGAGAGCTTTGAAGTCCAGAAAGGCT 180
Db 220 TCCATCAAAATCGAGATGCTGATGACACCCAGAGAGCTTTGAAGTCCAGAAAGGCT 279
QY 181 CCTGGAGAGGTGCTGCTGATGACAGTTTGCAACCACTCAACTCGTGAAGGTGACTAT 240
Db 280 CCTGGAGAGGTGCTGCTGATGACAGTTTGCAACCACTCAACTCGTGAAGGTGACTAT 339
QY 241 TTTGGCTCGAGTTTCTCTGATCAAAAAGATCAAGGTGCTGCTGATCTCTTAAACC 300
Db 340 TTTGGCTCGAGTTTCTCTGATCAAAAAGATCAAGGTGCTGCTGATCTCTTAAACC 399
QY 301 ATTGTGAACAGATTGGAAGGCCAAAGCAGTTGTTGTTAAGTTTGTGGAATTCCTT 360
Db 400 CTGTGTAACAGATTCAGAGGCCAAAGCAGTTGTTGTTAAGTTTGTGGAATTCCTT 459
QY 361 CCGCTGACACACACAACTCCAAGAACTCAAGAGTACCTGTTCGGCTGCAAGT 420
Db 460 CCGCTGACACACAACTCCAAGAACTCAAGAGTACCTGTTCGGCTGCAAGT 519
QY 421 AAGCAGGACTTGGCTCAAGCAGGTGACGTGTATGACACAGCCGACTCTCTTGATT 480
Db 520 AAGCAGGACTTGGCTCAAGCAGGTGACGTGTATGACACAGCCGACTCTCTTGATT 579
QY 481 TCACACATTTGCAATCTGATTTGGGGATTTTGTAGAAAGCTTTGACAGAGACCTTA 540
Db 580 TCACACATTTGCAATCTGATTTGGGGATTTTGTAGAAAGCTTTGACAGAGACCTTA 639
QY 541 GCAAAAATTAATACATCTCTGACAAAGCAGCTAGAGGCAAAAATCGTGAATTCAC 600
Db 640 GCAAAAATTAATACATCTCTGACAAAGCAGCTAGAGGCAAAAATCGTGAATTCAC 699
QY 601 CATTAACCATTTGACAAACACAGCAGAAATCAAGTTTCCAGTCTCTAGAGATTCGCCGT 660
Db 700 CATTAACCATTTGACAAACACAGCAGAAATCAAGTTTCCAGTCTCTAGAGATTCGCCGT 759
QY 661 CCGCTAGAGATGATGGAATCCGGTTGACCCGGCCAAAGACAGGAGGACAGAAATC 720
Db 760 CCGCTAGAGATGATGGAATCCGGTTGACCCGGCCAAAGACAGGAGGACAGAAATC 819
QY 721 AATCTGCGCTTCCAAACAGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
Db 820 AATCTGCGCTTCCAAACAGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 879
QY 781 TTCAATCGGGCCAAAGTGGGAAAGCTGAGTTCAAGAGAAAGGCTTTCTCATCAAGCTC 840
Db 880 TTCAATCGGGCCAAAGTGGGAAAGCTGAGTTCAAGAGAAAGGCTTTCTCATCAAGCTC 939
QY 841 CCGCAGATGCAATAGTGTGCTACAGAGATCACTTGAATTTCTGATGCGCAGT 900
Db 940 CCGCAGATGCAATAGTGTGCTACAGAGATCACTTGAATTTCTGATGCGCAGT 999
QY 901 TTCTGCAAGTCTTCTGGAATCTGTGTGAACATCAGCTCTTTAGACTTTTGA 960
Db 1000 TTCTGCAAGTCTTCTGGAATCTGTGTGAACATCAGCTCTTTAGACTTTTGA 1059
QY 961 GAGCCCAAAACCAAGCCCGCTCTTTAGCCGGGGTCACTTTCCGTTCAAGT 1020
Db 1060 GAGCCCAAAACCAAGCCCGCTCTTTAGCCGGGGTCACTTTCCGTTCAAGT 1119
QY 1021 GGTGCGACTCAGAGAGGTTCTGACTATGTTAAAGAGAGAGACATAAGAGTGCAG 1080
Db 1120 GGTGCGACTCAGAGAGGTTCTGACTATGTTAAAGAGAGAGACATAAGAGTGCAG 1179
QY 1081 TTGAAAGAGAGACAGCAAGATTCATTCTATCCGAGCCTTGCTTCAAGCTTACAGAA 1140

Db 1180 TTGAAAGAGAGACAGCAAGATTCATTCTATCCGAGCCTTGCTTACAGCTTACAGAA 1239
QY 1141 CTGAATTCGGAAGTGTGAGAGAGTCTACAGAGACAGACCTTACATTTGGAGAAAGT 1200
Db 1240 CTGTATTCGGAAGTGTGAGAGAGTCTACAGAGAGCCAGCTTACATTTGGAGAAAGT 1299
QY 1201 GCCGAATCTCAGAGGGGCGAGAGTGCCTGCGCAGAGAAAGAAACCGAAGGTTTCCGCGGG 1260
Db 1300 GCCGAATCTCAGAGGGGCGAGAGTGCCTGCGCAGAGAAAGAAACCGAAGGTTTCCGCGGG 1359
QY 1261 GAGCCGGGGTCCGACCCGAGCCCTGCGCAGAGAAAGCCCGCGGTAAACAAGCAGCG 1320
Db 1360 GAGCCGGGGTCCGACCCGAGCCCTGCGCAGAGAAAGCCCGCGGTAAACAAGCAGCG 1419
QY 1321 GAGGAGCCGCTCGGCGCCAGAGGAGAGAGAGAGAGTGTGTTAAGATTAGAACCA 1380
Db 1420 GAGGAGCCGCTCGGCGCCAGAGGAGAGAGAGAGTGTGTTAAGATTAGAACCA 1479
QY 1381 CAGAGTAACTCAGCCCGCAGCCAGCAGCAGAGCTCCCTGACTGCGAGTCCCTCACTT 1440
Db 1480 CAGAGTAACTCAGCCCGCAGCCAGCAGCAGAGCTCCCTGACTGCGAGTCCCTCACTT 1539
QY 1441 TCGAGCTGTCTGTGAATCTGCAAGGGGGAGTGGCCCTGCAACGTGACTTGTCTCC 1500
Db 1540 TCGAGCTGTCTGTGAATCTGCAAGGGGGAGTGGCCCTGCAACGTGACTTGTCTCC 1599
QY 1501 AACCTGAGCCCGCAGCAGCAGAGAGGCTCCCTTGATGAGCCGCTGATGACAG 1560
Db 1600 AACCTGAGCCCGCAGCAGCAGAGAGGCTCCCTTGATGAGCCGCTGATGACAG 1659
QY 1561 GCTGCCCCCGAGCAGCAGATGAGAGGCGCGAGAGAGATTCCTCAACTGATAA 1620
Db 1660 GCTGCCCCCGAGCAGCAGATGAGAGTGAAGGCGCGAGAGAGATTCCTCAACTGATAA 1719
QY 1621 GCGTACTTATAGCTAAGAGATGTCTTACACCGAGCCAACTATCTGAAGATCTCGAA 1680
Db 1720 GCGTACTTATAGCTAAGAGATGTCTTACACCGAGCCAACTATCTGAAGATCTCGAA 1779
QY 1681 GTTATCACTTGGTGGTTTCAAGACAGTGAACAAAGAGAGAGCCATGCGGAGCACTG 1740
Db 1780 GTTATCACTTGGTGGTTTCAAGACAGTGAACAAAGAGAGAGCCATGCGGAGCACTG 1839
QY 1741 AAAAGTCTCATATTCCTCGAATTTTGAACCTTTGCAAAATTTCTATATTTTCTCAAG 1800
Db 1840 AAAAGTCTCATATTCCTCGAATTTTGAACCTTTGCAAAATTTCTATATTTTCTCAAG 1899
QY 1801 GAAATGAGCAACGACTTGCCTGTGGGAAAGCCGCTCAAAATGCCAATTCAGAGTTAC 1860
Db 1900 GAAATGAGCAACGACTTGCCTGTGGGAAAGCCGCTCAAAATGCCAATTCAGAGTTAC 1959
QY 1861 CAAAGATCGGCGAGTCACTGCTGAAGAACTTCAAGGGCATGAAAGCACTGGCGGCTCAC 1920
Db 1960 CAAAGATCGGCGAGTCACTGCTGAAGAACTTCAAGGGCATGAAAGCACTGGCGGCTCAC 2019
QY 1921 CTGTGAGACACAGCGGCTTGGAGGCGCTCTGAGAAATGGAATCAAGACTCCCGCGG 1980
Db 2020 CTGTGAGACACAGCGGCTTGGAGGCGCTCTGAGAAATGGAATCAAGACTCCCGCGG 2079
QY 1981 CTGAGAACTTCTGCAAGAGACTTTGAGCTGCAAGAGTGTGTTACTTACCGCTCAACAC 2040
Db 2080 CTGAGAACTTCTGCAAGAGACTTTGAGCTGCAAGAGTGTGTTACTTACCGCTCAACAC 2139
QY 2041 TTCTCTCGCGGCACTGACCGGCTCANTGCACTCAAGCAGAGTCTCTGAGCGGCTGTG 2100
Db 2140 TTCTCTCGCGGCACTGACCGGCTCANTGCACTCAAGCAGAGTCTCTGAGCGGCTGTG 2199
QY 2101 AAACACACCCCGCAGACCGCAGACTTCAAGGACTGCGAGCCGCTTTGGCAGAGATC 2160
Db 2200 AAACACACCCCGCAGACCGCAGACTTCAAGGACTGCGAGCCGCTTTGGCAGAGATC 2259
QY 2161 ACGAGATGTGCAAGTCTCAAGGATGATGATCAAGTGAAGATTTCCAGAAAGCTG 2220

Db 2260 ACGAGATGTTGGTGCACAGCTCCAGGTACGATGATCAAGATGAGAAATTTCCAGAGCTG 2319
 Qy 2221 CACGAATTCAGAAAGATTTGATTTGGCAATCTTGTGTTCCGGGAAAGGAGTTC 2280
 Db 2320 CACGAATTCAGAAAGATTTGATTTGGCAATCTTGTGTTCCGGGAAAGGAGTTC 2319
 Qy 2281 ATCCGTCTGGGACAGCTCAGCAAGCTCTCGGGAAAGGAGTTCAGCAAGCTTTCTTC 2340
 Db 2380 ATCCGTCTGGGACAGCTCAGCAAGCTCTCGGGAAAGGAGTTCAGCAAGCTTTCTTC 2439
 Qy 2341 CTGTTCACGACGCTCTGTATACGACGAGCCGGGGCTGACGCGCTCCATCACTTTAAA 2400
 Db 2440 CTGTTCACGACGCTCTGTATACGACGAGCCGGGGCTGACGCGCTCCATCACTTTAAA 2499
 Qy 2401 GTCCACGGGAGCTCCCGCTCTATAGCATGACATTTAGAGAGACGAGAGTGGGGG 2460
 Db 2500 GTCCACGGGAGCTCCCGCTCTATAGCATGACATTTAGAGAGACGAGAGTGGGGG 2559
 Qy 2461 GTGCCCCAGTGCCTGACCTCCGGGGCAGCGGAGTCCATCATGTCGCGCAGTTCT 2520
 Db 2560 GTGCCCCAGTGCCTGACCTCCGGGGCAGCGGAGTCCATCATGTCGCGCAGTTCT 2619
 Qy 2521 CGGTCCGAGATGAGAAAGTGGTTGAGGACATCCAGATGGCCATTTGACTGGCGAGAG 2580
 Db 2620 CGGTCCGAGATGAGAAAGTGGTTGAGGACATCCAGATGGCCATTTGACTGGCGAGAG 2679
 Qy 2581 AGCAGACGCCCCCGCTGAGTTCTTGGCCAGCAGCCCCCTGACAAAGTCCCTGAT 2640
 Db 2680 AACAGACGCTCCCGCTGAGTTCTTGGCCAGCAGCCCCCTGACAAAGTCCCTGAT 2739
 Qy 2641 GAAGCCACGCGGCTGACAGAGATGAGAGATGAGAGCTGAGCGCTCGCGACATCGCTG 2700
 Db 2740 GAAGCCACGCGGCTGACAGAGATGAGAGATGAGAGCTGAGCGCTCGCGACATCGCTG 2799
 Qy 2701 GAGCGCCAGGCCCCCGACCGCGGACACAAATGTTGCTGCGACACCGCACACC 2760
 Db 2800 GAGCGCCAGGCCCCCGACCGGTGGACAAATGTTGATGTTCTGGACACCGCACACC 2859
 Qy 2761 AGGTCTCCATGTTGAGCTTCAAGATGCGAGTGGAGAAATCAGTTGTCTGAAAACCTGCTG 2820
 Db 2860 AGGTCTCCATGTTGAGCTTCAAGATGCGAGTGGAGAAATCAGTTGTCTGAAAACCTGCTG 2919
 Qy 2821 AGGAATTTCAAAAACAGCAACGAGTGGAGAGTGTGGTGTCTCAAAACCTTGTG 2880
 Db 2920 AGGAATTTCAAAAACAGCAACGAGTGGAGAGTGTGGTGTCTCAAAACCTTGTG 2979
 Qy 2881 CTGTCTCTTCAAAATCAACAGCAATCATCCCTTGCAGCTGCTGCTGCGC 2940
 Db 2980 CTGTCTCTTCAAAATCAACAGCAATCATCCCTTGCAGCTGCTGCTGCGC 3039
 Qy 2941 TACTCGCTCACTCCCTCTGAGTCCGAGAACATCCAGAAAGCTAGTGTTCAGCTG 3000
 Db 3040 TACTCGCTCACTCCCTCTGAGTCCGAGAACATCCAGAAAGCTAGTGTTCAGCTG 3099
 Qy 3001 CACTTCAAGTCCAGCTCTACTCTTCAAGGCGGAAAGAGTACAGTTGCAAGGTG 3060
 Db 3100 CACTTCAAGTCCAGCTCTACTCTTCAAGGCGGAAAGAGTACAGTTGCAAGGTG 3159
 Qy 3061 ATGGAATGATCCGACAGTGCACCAAGCTTGCCTGCGACCCCACTGTTGAGCCAAA 3120
 Db 3160 ATGGAATGATCCGACAGTGCACCAAGCTTGCCTGCGACCCCACTGTTGAGCCAAA 3219
 Qy 3121 GAGTCTCTTGTGAT 3135
 Db 3220 GAGTCTCTTGTGAT 3234

RESULT 3
 AY406168 3138 bp DNA linear GSS 15-DEC-2003
 LOCUS Pan troglodytes FAPR1 gene, VIRUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 AY406168
 ACCESSION

VERSION AY406168.1 GI:39762142
 KEYWORDS GSS
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Mammalia; Eutelestomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi; Mammalia; Eutelestia; Euarchontoglires; Primates; Catarrhini; Homidae; Pan.
 REFERENCE 1 (bases 1 to 3138)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Smitsky,J.J., Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 3138)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Smitsky,J.J., Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment
 FEATURES
 source location/Qualifiers
 1..3138
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>3138
 /gene="FAPR1"
 /locus_tag="HCW2467"
 ORIGIN
 Query Match 94.6%; Score 2965.2; DB 14; Length 3138;
 Best Local Similarity 94.8%; Pred. No. 0;
 Matches 2973; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
 Qy 1 ATGGAGAAATAGAGCAGAGCCGACCCAGATCAAGTGGGGGCCGGAATTCG 60
 Db 1 ATGGAGAAATAGAGCAGAGCCGACCCAGATCAAGTGGGGGCCGGAATTCG 60
 Qy 61 GGGATCATGTCCTTGGAACTGAGCAGAGCCGCCCAACCTTCAGAAATCTGTG 120
 Db 61 GGGATCATGTCCTTGGAACTGAGCAGAGCCGCCCAACCTTCAGAAATCTGTG 120
 Qy 121 TCCATCAAAATCCAGATGCTGATGACACCAGAGGCAATTTGAAGTCCAAAAGACT 180
 Db 121 TCCATCAAAATCCAGATGCTGATGACACCAGAGGCAATTTGAAGTCCAAAAGACT 180
 Qy 181 CTTGGAGAGGTGCTGTGATGATGCACTTCAACCTCGTGAAGGTGACTAT 240
 Db 181 CTTGGAGAGGTGCTGTGATGATGCACTTCAACCTCGTGAAGGTGACTAT 240
 Qy 241 TTGGGCTCGAGTTCTCTGATCAAAAAGATACAGGTGCTGATCTCTPAAACC 300
 Db 241 TTGGGCTCGAGTTCTCTGATCAAAAAGATACAGGTGCTGATCTCTPAAACC 300
 Qy 241 NNN 300
 Db 241 NNN 300
 Qy 301 ATTGTGAACAGATTGAAGGCCAAGACGTTGTTGAAGTTTGTGGAATCTTT 360
 Db 301 ATTGTGAACAGATTGAAGGCCAAGACGTTGTTGAAGTTTGTGGAATCTTT 360
 Qy 361 CCGCTGACCAACACATCCCAAGAAAGCTCACAAGGTACTGTTGGCTGACAGTG 420
 Db 361 CCGCTGACCAACACATCCCAAGAAAGCTCACAAGGTACTGTTGGCTGACAGTG 420
 Qy 421 AACGAGACTTGGCTCAAGGCAAGTGTGATGACACCGAGCCAGCTCTTGATT 480
 Db 421 AACGAGACTTGGCTCAAGGCAAGTGTGATGACACCGAGCCAGCTCTTGATT 480
 Qy 481 TCACACTTGTGAACTGTGATTTGGGATTTTGTGTAAGACTTGAAGAGACCTTA 540

Db 481 TCACATCTGTGCAATCTGAGATTTGGGGATTTTGTAGTAAGCTTTGGACAGAGGACCTTA 540
QY 541 GCAAAAAATTAATATCTACTCTGACAAAGACCTAGAGGACAAATCGTGAATTTCCAC 600
Db 541 GCAAAAAATTAATATCTACTCTGACAAAGACCTAGAGGACAAATCGTGAATTTCCAC 600
QY 601 CATTAACCATTTGGAGCAAAACACAGCAGATCAGATTTCCAGCTCTAGAGATTTGCCGT 660
Db 601 CATTAACCATTTGGAGCAAAACACAGCAGATCAGATTTCCAGCTCTAGAGATTTGCCGT 660
QY 661 CGGCTAGAGATGATGGAATCCGGTTGCAACCGGCAAGGACAGGAGGCAAGAAATC 720
Db 661 CGGCTAGAGATGATGGAATCCGGTTGCAACCGGCAAGGACAGGAGGCAAGAAATC 720
QY 721 AATCTGGCCGTTGGCCAAACAGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
Db 721 AATCTGGCCGTTGGCCAAACAGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
QY 781 TTCACTGGGCGCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGGCTTTCTCATCAAGCTC 840
Db 781 TTCACTGGGCGCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGGCTTTCTCATCAAGCTC 840
QY 841 CGGCAAGATGCAATAGTGGTACAGAGATACCTTGAATTCCTGATGGCAGTGGGAT 900
Db 841 CGGCAAGATGCAATAGTGGTACAGAGATACCTTGAATTCCTGATGGCAGTGGGAT 900
QY 901 TTCTGCAAGCTCTTCTGGAATAATCTGTGTGAACATCAATGCTTTTGAATCTTTTGA 960
Db 901 TTCTGCAAGCTCTTCTGGAATAATCTGTGTGAACATCAATGCTTTTGAATCTTTTGA 960
QY 961 GAGCCCAACCAAGCCCAAGCCCGTCTTTCAGCGGGGTCATCATTTCCGTTCACT 1020
Db 961 GAGCCCAACCAAGCCCAAGCCCGTCTTTCAGCGGGGTCATCATTTCCGTTCACT 1020
QY 1021 GGTGGAATCTGAGAGAGGTTCTGACTATGTTAAAGAGAGACATAAGAGGTGCAG 1080
Db 1021 GGTGGAATCTGAGAGAGGTTCTGACTATGTTAAAGAGAGACATAAGAGGTGCAG 1080
QY 1081 TTTGAAAGAGAGACAGCAAGATTTCAATTCATTCGAGACCTTGCTTCAAGCCTTACAGAA 1140
Db 1081 TTTGAAAGAGAGACAGCAAGATTTCAATTCATTCGAGACCTTGCTTCAAGCCTTACAGAA 1140
QY 1141 CTGAATTCGGAAGTGTGAGCAGTCTCAGCAGAGACACAGCCTTACATTTGAGAGAGT 1200
Db 1141 CTGAATTCGGAAGTGTGAGCAGTCTCAGCAGAGACACAGCCTTACATTTGAGAGAGT 1200
QY 1201 GCCGAATCTTCAGAGGGGCCAGAGCTGCGGCGAGAGAAAGAACCGAAGGTTTCCGCCGG 1260
Db 1201 GCCGAATCTTCAGAGGGGCCAGAGCTGCGGCGAGAGAAAGAACCGAAGGTTTCCGCCGG 1260
QY 1261 GAGCCGGGGTGCAGACCGAGCCCTGCGCGAGAGAAAGCCCGGGGTAAACAAGAGGG 1320
Db 1261 GAGCCGGGGTGCAGACCGAGCCCTGCGCGAGAGAAAGCCCGGGGTAAACAAGAGGG 1320
QY 1321 GACGAGACCGCTCGGCGGCCACGAGAGAGAGAGAGAGTCTGTTAAGATAGAACCCAG 1380
Db 1321 GACGAGACCGCTCGGCGGCCACGAGAGAGAGAGAGAGTCTGTTAAGATAGAACCCAG 1380
QY 1381 CAGAGTAACTCAGCCCCCGCAGACCAAGACAGAGCTCCCTGACTGGCAGTCTTCACTT 1440
Db 1381 CAGAGTAACTCAGCCCCCGCAGACCAAGACAGAGCTCCCTGACTGGCAGTCTTCACTT 1440
QY 1441 TCCGACCTGTCTGTGAATCTGCAAGGGGGAGTGGCCCTGSCCAACGTAACCTTGTCTCC 1500
Db 1441 TCCGACCTGTCTGTGAATCTGCAAGGGGGAGTGGCCCTGSCCAACGTAACCTTGTCTCC 1500
QY 1501 AACCTGAGCCCGACACCAAGCAGGCTCTCCCTTGATCAGCCCGCTGTGAATGACAG 1560
Db 1501 AACCTGAGCCCGACACCAAGCAGGCTCTCCCTTGATCAGCCCGCTGTGAATGACAG 1560
QY 1561 GCCTGCCCCCGACGAGATGAGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATAAA 1620

Db 1561 GCCTGCCCCCGACGAGCAGTGAAGTGAAGGCCCGGAGAGAGATTTCCCACTGATAAA 1620
QY 1621 GGGTACTTATAGCTAAGGAAGTGTCTACACCGAGCGAACTATCTGAAGATCTCGAA 1680
Db 1621 GGGTACTTATAGCTAAGGAAGTGTCTACACCGAGCGAACTATCTGAAGATCTINGAA 1680
QY 1681 GTTATCACTTGTGTGTTTCAAGACACAGTGAACAAAGAGACCCATGCGGAGCAGT 1740
Db 1681 GTTATCACTTGTGTGTTTCAAGACACAGTGAACAAAGAGACCCATGCGGAGCAGT 1740
QY 1741 AAAAGTCTATATTTCCGGAATTTTGAACCTTTGACAAATTTCTATCTAATTTTCAAG 1800
Db 1741 AAAAGTCTATATTTCCGGAATTTTGAACCTTTGACAAATTTCTATCTAATTTTCAAG 1800
QY 1801 GAAATTAAGCAAGACTTCCCTGTGGAAGGCGGCTCAATGSCCAATTCAGAGATAC 1860
Db 1801 GAAATTAAGCAAGACTTCCCTGTGGAAGGCGGCTCAATGSCCAATTCAGAGATAC 1860
QY 1861 CAAAGATCGGGAGATGTCTGTGAAGAACATTCAGGGCATGAAGCCTGGCGGCTCAC 1920
Db 1861 CAAAGATCGGGAGATGTCTGTGAAGAACATTCAGGGCATGAAGCCTGGCGGCTCAC 1920
QY 1921 CTGTGGAACACAGAGAGCTTTGAGAGCCCTGAGAGAAATGAATCAAGAGCTTCCGCGG 1980
Db 1921 CTGTGGAACACAGAGAGCTTTGAGAGCCCTGAGAGAAATGAATCAAGAGCTTCCGCGG 1980
QY 1981 CTGGAAGACTTCTGAGAGACTTTGAGCTGAGAGAGGTGTGTTACTTACCGCTCAACACC 2040
Db 1981 CTGGAAGACTTCTGAGAGACTTTGAGCTGAGAGAGGTGTGTTACTTACCGCTCAACACC 2040
QY 2041 TTCTTCTGCGGCACTGACCGGCTCATGCACTACAGCAGTCTCTGAGCGGCTGTGC 2100
Db 2041 TTCTTCTGCGGCACTGACCGGCTCATGCACTACAGCAGTCTCTGAGCGGCTGTGC 2100
QY 2101 AAAACACACCCCGGAGCCAGCCGACTTCAAGGACTGCGGACCGCTTTGGCAGAGATC 2160
Db 2101 AAAACACACCCCGGAGCCAGCCGACTTCAAGGACTGCGGACCGCTTTGGCAGAGATC 2160
QY 2161 ACGAGATGATGAGCAGCTCCACGGTACGATGATCAAGATGAGAAATTTCCAGAGCTG 2220
Db 2161 ACGAGATGATGAGCAGCTCCACGGTACGATGATCAAGATGAGAAATTTCCAGAGCTG 2220
QY 2221 CACGAATCTAAGAAAGATTTGATTGGCATTGACAACTTGTGGTCCGGAAAGGAGTTC 2280
Db 2221 CACGAATCTAAGAAAGATTTGATTGGCATTGACAACTTGTGGTCCGGAAAGGAGTTC 2280
QY 2281 ATCCGTCTGGGAGCTTCAGCAAGCTCTCGGGAGAGGGCTCCAGAGGCGATGTTCTTC 2340
Db 2281 ATCCGTCTGGGAGCTTCAGCAAGCTCTCGGGAGAGGGCTCCAGAGGCGATGTTCTTC 2340
QY 2341 CTGTTCAACGAGCTCTGCTATACAGAGCCGGGGCTGACGGCTCCATCAGTTTAAA 2400
Db 2341 CTGTTCAACGAGCTCTGCTATACAGAGCCGGGGCTGACGGCTCCATCAGTTTAAA 2400
QY 2401 GTTCACGGGCACTCCCGCTATGAGCATTGAGAGACGAGACGAGTGGGG 2460
Db 2401 GTTCACGGGCACTCCCGCTATGAGCATTGAGAGACGAGACGAGTGGGG 2460
QY 2461 GTGCCCACTGCTGACCTTCGGGGCCAGCGGCGAGTCCATATGCTGGCCCGCAGTTCT 2520
Db 2461 GTGCCCACTGCTGACCTTCGGGGCCAGCGGCGAGTCCATATGCTGGCCCGCAGTTCT 2520
QY 2521 CGGTCCGAGATGAGAGAGTGGTTGAGGACATCCAGATGGCATTTGACCTGGCGAGAA 2580
Db 2521 CGGTCCGAGATGAGAGAGTGGTTGAGGACATCCAGATGGCATTTGACCTGGCGAGAA 2580
QY 2581 AGCAGAGCCCGCCCTTGAGTTCTGTGCGCAGACGCCCCCTGACAAAGTCCCTGAT 2640
Db 2581 AGCAGAGCCCGCCCTTGAGTTCTGTGCGCAGACGCCCCCTGACAAAGTCCCTGAT 2640
QY 2641 GAAAGCCACCGCGCTGACCAAGAGTCAAGAGATGACCTGAGCGCTCGCGCACGTCGTG 2700
Db 2641 GAAAGCCACCGCGCTGACCAAGAGTCAAGAGATGACCTGAGCGCTCGCGCACGTCGTG 2700

QY 2701 GAGCGCCAGGCCCCGCAACCGCGCAACAATGTGACGCTGTGGACCCGCAACCC 2760
 |||||
 Db 2701 GAGCGCCAGGCCCCGCAACCGCGCAACAATGTGACGCTGTGGACCCGCAACCC 2760
 |||||
 QY 2761 AGCGCTCCATGTGTGACCTTACGATGCGAGTGAATCATGTTGTGTGAAACCTGCTG 2820
 |||||
 Db 2761 AGCGCTCCATGTGTGACCTTACGATGCGAGTGAATCATGTTGTGTGAAACCTGCTG 2820
 |||||
 QY 2821 AGGAAATTCAAAACACAAACCGGTGCGAAGAGTGTGGTGTTCACAAATTTCTGC 2880
 |||||
 Db 2821 AGGAAATTCAAAACACAAACCGGTGCGAAGAGTGTGGTGTTCACAAATTTCTGC 2880
 |||||
 QY 2881 CTGTCTTCTTACAATATACACAGAGATCATCCCTTGCAGACCTGCTCTGCTGCGC 2940
 |||||
 Db 2881 CTGTCTTCTTACAATATACACAGAGATCATCCCTTGCAGACCTGCTCTGCTGCGC 2940
 |||||
 QY 2941 TACTCGCTACCATCCCTCTGTGAGTCCGAGAACATCCGAAAGACTACGTGTCAAGCTG 3000
 |||||
 Db 2941 TACTCGCTACCATCCCTCTGTGAGTCCGAGAACATCCGAAAGACTACGTGTCAAGCTG 3000
 |||||
 QY 3001 CACTTCAAGTCCCAACGCTTACTTACTTCAAGGCGGAAAGCGATCACGTTGCAAAAGTGG 3060
 |||||
 Db 3001 CACTTCAAGTCCCAACGCTTACTTACTTCAAGGCGGAAAGCGATCACGTTGCAAAAGTGG 3060
 |||||
 QY 3061 ATGGAATGATCCGACAGTGCACACAGCTCTGCTCGGACCCGACGTTGAGCCACAA 3120
 |||||
 Db 3061 ATGGAATGATCCGACAGTGCACACAGCTCTGCTCGGACCCGACGTTGAGCCACAA 3120
 |||||
 QY 3121 GAGTCTCTTGTGTAT 3135
 |||||
 Db 3121 GAGTCTCTTGTGTAT 3135
 |||||

RESULT 4
 AY406169 3147 bp DNA linear GSS 15-DEC-2003
 LOCUS Mus musculus FARP1 gene, VIRUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY406169
 VERSION AY406169.1 GI:39762143
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 1 (bases 1 to 3147)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civejlo,D.R., Lu,F., Murphy,B.,
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Snieksy,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene clusters
 Science 302 (5652), 1960-1963 (2003)

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 2 (bases 1 to 3147)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civejlo,D.R., Lu,F., Murphy,B.,
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Snieksy,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submision
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 COMMENT
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 1. 3147
 Location/Qualifiers
 source
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1. >3147
 /gene="FARP1"

ORIGIN
 /Locus_tag="HGM2467"

Query Match 77.4%; Score 2425; DB 14; Length 3147;
 Best Local Similarly 86.2%; Pred. No. 0;
 Matches 2709; Conservative 0; Mismatches 425; Indels 9; Gaps 2;

QY 1 ATGGAGAAATATAGCAGAGAGCCGACCGAGATCAGATCGGGGGCCCCGAAATTCG 60
 |||||
 Db 1 ATGGAGAAATATAGCAGAGAGCCGACCGAGATCAGATCGGGGGCCCCGAGAACTCG 60
 |||||
 QY 61 GGGATCAGTACTTGGAAACGTGACAGAGCCGCCCAACACCTTCAGAGAAATCTGTC 120
 |||||
 Db 61 GGGATCAGTACTTGGAAACGTGACAGAGCCGCCCAACACCTTCAGAGAAATCTATG 120
 |||||
 QY 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAGTTCCAAAGACT 180
 |||||
 Db 121 ACTGTCAAAATTCAGATGCTGATGACACCCAGAGAGCATTTGAAGTTCCAAAGACT 180
 |||||
 QY 181 CTTGGGAAGGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 |||||
 Db 181 CTTGGGAAGGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 |||||
 QY 241 TTTGGCTCGAGTTTCTGATCAAAAGATCAGCTGTGCTGATCTCTAAAACCC 300
 |||||
 Db 241 TTTGGCTCGAGTTTCTGATCAAAAGATCAGCTGTGCTGATCTCTAAAACCC 300
 |||||
 QY 301 ATTGTGAACAGATTGAGAGGCCAAGACAGCTGTGTTGTTAAGTTGTGTTGTTT 360
 |||||
 Db 301 ATTGTGAACAGATTGAGAGGCCAAGACAGCTGTGTTGTTAAGTTGTGTTGTTT 360
 |||||
 QY 361 CCGCTGACACACACACACCTCCAAAGAACTCAAGATACCTGTTGCTGCGAGGTG 420
 |||||
 Db 361 CCGCTGACACACACACACCTCCAAAGAACTCAAGATACCTGTTGCTGCGAGGTG 420
 |||||
 QY 421 AACGAGACTTGGCTCAAGCAGTTCAGCTGATATGACACAGCAGCTCTTGTATT 480
 |||||
 Db 421 AACGAGACTTGGCTCAAGCAGTTCAGCTGATATGACACAGCAGCTCTTGTATT 480
 |||||
 QY 481 TCACACATTTGTGAATCTGATGATTTGGGATTTGATGAAAGCTTGGACAGAGCACTTA 540
 |||||
 Db 481 TCACACATTTGTGAATCTGATGATTTGGGATTTGATGAAAGCTTGGACAGAGCACTTA 540
 |||||
 QY 541 GCAAAATTAATTAACATCTCTGCAAGACGCACTGAGAGCAAAATTCGGAATTCAC 600
 |||||
 Db 541 GCAAAATTAATTAACATCTCTGCAAGACGCACTGAGAGCAAAATTCAGAGTTCCAC 600
 |||||
 QY 601 CATTAACCATTTGAGCAAAACACAGCAGATCATGTTTCCAGCTCTAGAGATTGCCGT 660
 |||||
 Db 601 CATTAACCATTTGAGCAAAACACAGCAGATCATGTTTCCAGCTCTAGAGATTGCCGT 660
 |||||
 QY 661 CGGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 |||||
 Db 661 CGGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 |||||
 QY 721 AATCTGCGCTGGCCCAACCGGAATTTAGTGTTCAGAGTTTCAAGTTCATCAATGCC 780
 |||||
 Db 721 AATCTGCGCTGGCCCAACCGGAATTTAGTGTTCAGAGTTTCAAGTTCATCAATGCC 780
 |||||
 QY 781 TTCAATGGGCGCAAGGTGCGAAAGTGAAGTTCAGAGAGAGGCTTTCATCAAGCTC 840
 |||||
 Db 781 TTCAATGGGCGCAAGGTGCGAAAGTGAAGTTCAGAGAGAGGCTTTCATCAAGCTC 840
 |||||
 QY 841 CGGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 |||||
 Db 841 CGGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 |||||
 QY 901 TTCTGAAGTCTTCTGAGAAATCTGTGTTGAACATCATGACCTTCTTGAAGTTTGA 960
 |||||
 Db 901 TTCTGAAGTCTTCTGAGAAATCTGTGTTGAACATCATGACCTTCTTGAAGTTTGA 960
 |||||
 QY 961 GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGTCATCATTTGGTTCACT 1020
 |||||

Dh 961 GAACCTTAAGCAAAAGCCAAAGCTGTTCTTCAGCCGAGGAGTGTCCTTCGAGTTAAGT 1020
Qy 1021 GGTGGAGCTCAAGACAGGTTCTCGACTATGTTTAAAGAGAGACATAAGAGGTGCAAG 1080
Db 1021 GGTCGAGCTCAAGAGCAAGTTCTTGATTTAAGTAAGAGCGCCACAAGAAAGGTGCAAG 1080
Qy 1081 TTTGAAGAGAGACAGCAAGATTCTATTCGAGCCTTGCTTCAAGCCTTAAGAA 1140
Db 1081 TTTGAAGAGAGACAGCAAGATCTCACTCAAGAGAGCTTGATCTCAAGCCCAAGCA 1140
Qy 1141 CTGAATTCGGAAGTGTGAGCAGTCTCAGCAGAGACAAGCCTTACATTTGAGAAAGT 1200
Db 1141 CCGAATTCAGAAATGCAAAACAGTCTCCACAGAGCGCAAGCCTTACGTTTGAGAAAGT 1200
Qy 1201 GCCGAATCTTCAGAGGGGCCAGAGCTGCGGCGAGAGAAAGAACCGAAAGTTCCGCGGG 1260
Db 1201 ACCGAGTCTTCGAGGTGCGAGAGCTGCGAGCAAGAGAAACAAAGCTGTATACCTTGG 1260
Qy 1261 GAGCCGAGGTCGACCCGAGCCCTGCGCGAGAGAGAGCCCGGAGTAAACAAGCAGGCG 1320
Db 1261 GAGTTGGAGCTTCAACAAAGCCCCGCTCTCTAAGAGCCCTCAGAGAGCAAAAGCTGCA 1320
Qy 1321 GACGAGACGCGCTCGGCGGCCACGAGAGAA-----GAGAGAGAGTCTGTTAAGATAG 1374
Db 1321 GATGGAACCAACGATGAGCCGCTGAGAGAAAGAGAGAGAGAGAGAGGTGCAAGACGCG 1380
Qy 1375 ACCGACAGATTAACCTCAGCCCCCGGAGACCAAGCAGAGCTCCTGACTGGCAATCCT 1434
Db 1381 ATCCGACCAAGTAAACCTCAGCCCCCGAGCCAGAGCAGAGCTCCTGATGATACCTCT 1440
Qy 1435 CACCTTTCGAGCTGTCTGTGAATCGAGAGGGGAGAGTGGCCCTGCAACGTCGACTTG 1494
Db 1441 CACCTTTCAGAGCTGTCTGTGAATCGAGAGGGGAGAGAGAGAGAGAGAGAGAGAG 1500
Qy 1495 TCTCCCAACCTGAGCCCGACACCAAGCAGAGCCTTCTCCTTATCAGCCGCTGTGAAT 1554
Db 1501 TCTCCCAACCTGAGCTGTCTGTGAATCAAGCAGAGCCTTCTCCTTATCAGCCGCTGTGAAT 1560
Qy 1555 GACCAAGGCTGCCCCCGAGAGAGAGATGAGAGATGAGAGGCGGAGAGAGAGATTTCCCACT 1614
Db 1561 GACCAAGGCTGCCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTCCCACT 1620
Qy 1615 GATTAAGAGCTTCTTATGAGAAAGTGTCTACCAACGAGCAACATCTGAAAGAT 1674
Db 1621 GATTAAGAGCTTCTTATGAGAAAGTGTCTACCAACGAGCAACATCTGAAAGAT 1680
Qy 1675 CTGAAGTATCACTTCGTGTTTCAAGACAGTGAAGCAAGAGAGAGAGAGAGAGAGAG 1734
Db 1681 CTGAAGTATCACTTCGTGTTTCAAGACAGTGAAGCAAGAGAGAGAGAGAGAGAGAG 1740
Qy 1735 GCACTGAAAGTCTCATATTCGGAATTTGAACTTTGCAAAATTTCTACTAATTTT 1794
Db 1741 GCTTTAAGAGTCTCATATTCGGAATTTGAACTTTGCAAAATTTCTACTAATTTT 1800
Qy 1795 CTGAAGAAATGAGCAAGACTTGCCTGTGGAAGAGCGCTCAATGCGCAATCA-- 1852
Db 1801 CTGAAGAAATGAGCAAGACTTGCCTGTGGAAGAGCGCGCTCAATGCGCAATCAAG 1860
Qy 1853 -GAGATTACCAAGAAATCGGAGATGTCTATGTAAGAAATTCAGGAGCATGAAGACTTG 1911
Db 1861 GAGATTACCAAGAAATCGGAGATGTCTATGTAAGAAATTCAGGAGCATGAAGACTTG 1920
Qy 1912 GCGGCTCACTGTGGAAGACAGCAGAGAGCTTGAAGGCTTGAAGATGGAATCAAGAGC 1971
Db 1921 GCGGCTCACTGTGGAAGACAGCAGAGAGCTTGAAGGCTTGAAGATGGAATCAAGAGC 1980
Qy 1972 TCCGAGGCTGAGAACTTCTGCAAGACTTGAAGAGTGTGTTACTTACCG 2031
Db 1981 TCCGAGGCTGAGAACTTCTCTGCAAGACTTGAAGAGTGTGTTACTTACCG 2040
Qy 2032 CTGAACACCTTCTCTGCGAGCACTGACACGAGCTCATGAACAGAGGCTCTGAG 2091
Db 2041 CTGAACACCTTCTCTGCGAGCACTGACACGAGCTCATGAACAGAGGCTCTGAG 2100

Qy 2092 CGGCTGTGCAACACACCCGCGAGGCCAGCCGACTTCAGAGGACTGCGGAGCCGCTTG 2151
Db 2101 AGCTGTGCAACACACCCGCGAGGCCAGCCGACTTCAGAGGACTGCGGAGCCGCTTG 2160
Qy 2152 GCGAGATCAACGAGATGTGAGCAGCTCAACGATGATCAAGATGAGAAATTTTC 2211
Db 2161 GCGAGATCAACGAGATGTGAGCAGCTGATGATGATCAAGATGAGAAATTTTC 2220
Qy 2212 CAGAACTGCAAGAACTCAAGAAATTTGATTTGCAATTAATTTGATTTGATTTG 2271
Db 2221 CAGAACTGCAAGAACTCAAGAAATTTGATTTGCAATTAATTTGATTTGATTTG 2280
Qy 2272 AGGAGTCACTCGTCTGAGGAGCCCTCAGCAAGCTCTGAGGAGAGAGGCTCCAGCAGCC 2331
Db 2281 AGGAGTCACTCGTCTGAGGAGCCCTCAGCAAGCTCTGAGGAGAGAGGCTCCAGCAGCC 2340
Qy 2332 ATGTTCTTCTGTTCAAGACGTCCTGTTATACAGAGCCGAGGAGCTGACGAGCTTCAAT 2391
Db 2341 ATGTTCTTCTGTTCAAGATGTTCTTGTGTTATCAGAGCCGAGGAGCTGACGAGCTTCAAT 2400
Qy 2392 CAGTTTAAAGTCAAGGAGAGCTCCCGCTTATGAGATGACATTTGAGAGAGAGAGC 2451
Db 2401 CAGTTTAAAGTCAAGGAGAGAGCTCCCACTATGAGATGACATTTGAGAGAGAGAG 2460
Qy 2452 GAGTGGGAGGTCGCCCACTGCTGACCTCGGAGGAGAGGAGAGAGCTCATTCGAGGCG 2511
Db 2461 GAGTGGGAGGTCGCCCACTGCTGAGCTCGGAGGAGAGGAGAGAGCTCATTCGAGGCG 2520
Qy 2512 GCCAGTTCCTGATCCGAGATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2571
Db 2521 GCCAGTTCCTGATCCGAGATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2580
Qy 2572 GGGAG 2631
Db 2581 GGGAG 2640
Qy 2632 TCCCTGATGAG 2691
Db 2641 TCCCTGATGAG 2700
Qy 2692 AATGCTGAG 2751
Db 2701 AATGCTGAG 2760
Qy 2752 CGCAACACAG 2811
Db 2761 CGCAACACAG 2820
Qy 2812 AACCTGCTGAG 2871
Db 2821 AACCTGCTGAG 2880
Qy 2872 AACCTGCTGAG 2931
Db 2881 AACCTGCTGAG 2940
Qy 2932 CTGCTGAGTACTGCTGACATTCCTCTGAGTCCGAGAAATTCAGAAAGAGAGAGAG 2991
Db 2941 CTGCTGAGTACTGCTGACATTCCTCTGAGTCCGAGAAATTCAGAAAGAGAGAGAG 3000
Qy 2992 TTTCAAGCTGCACTTCAAGTCCGAGAGCTTCACTTCAAGGAGAGAGAGAGAGAGAG 3051
Db 3001 TTTCAAGCTGCACTTCAAGTCCGAGAGCTTCACTTCAAGGAGAGAGAGAGAGAGAG 3060
Qy 3052 GAAAGTGAATGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3111
Db 3061 GAAAGTGAATGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
Qy 3112 AGCACAAGAGTCTTCTGTTA 3134
Db 3121 AGCACAAGAGTCTTCTGTTA 3143

- RESULT 5
AK171713 3743 bp mRNA linear HTC 21-SEP-2005
- LOCUS
AK171713
- DEFINITION
Mus musculus activated spleen cDNA, RIKEN full-length enriched library, clone:830006D05 product:FSRM, Rhogef and pleckstrin domain protein 2, full insert sequence.
- ACCESSION
AK171713
- VERSION
AK171713.1 GI:74152707
- KEYWORDS
HTC; CAP trapper.
- SOURCE
Mus musculus (house mouse)
- ORGANISM
Mus musculus
- REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
- REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
- JOURNAL
PUBMED
11042159
- REFERENCE
3
Shibata, K., Itoh, M., Alzawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubara, M., Ozawa, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
- JOURNAL
PUBMED
11076861
- REFERENCE
4
Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Alzawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staab, J. F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Mammucari, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Wittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S., and Hayashizaki, Y.
Riken Genome Exploration Research Group Phase II Team and the FANTOM Consortium
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
- JOURNAL
PUBMED
11217851
- REFERENCE
5
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oseko, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanepin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Bruschi, V., Chotula, C., Corbani, L. E., Cousins, S., Dalla, E., Dragan, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kashiwagi, K., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numa, K., Okido, T., Pavan, W. J., Pettes, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Vitaro, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-kenikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Alzawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E., and Hayashizaki, Y.
PANTOM Consortium
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420 (6915), 563-573 (2002)
- JOURNAL
PUBMED
12468851
- REFERENCE
6
Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V. B., Brenner, S. E., Batalov, S., Forrest, A. R., Zavolan, M., Davis, M. J., Wilming, L. G., Aldrich, V., Allen, J. E., Ambesi-Imbulceto, A., Apweiler, R., Attalaya, R. N., Bailey, T. L., Bansal, M., Baxter, J., Beisel, K. W., Bersano, T., Bono, H., Chalk, A. M., Chiu, K. P., Choudhary, V., Christoffels, A., Cluttenbuck, D. R., Crome, M. L., Dalla, E., Dalrymple, B. P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C. F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T. R., Gojobori, T., Green, R. E., Gustincich, S., Haberer, M., Hayashi, Y., Hensch, T. K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kello, J., Kitamura, H., Kitano, H., Kohlhaas, G., Krishnan, S. P., Kruger, A., Kummerfeld, S. K., Kurochkin, I. V., Larau, L. F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan Babu, M., Mader, M., Marchionni, L., Matsuda, H., Matsunawa, S., Miki, H., Mignone, F., Miyake, S., Moris, K., Mortazavi, T., Muzier, S., Mulder, N., Nakano, N., Nakaguchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K. C., Pavan, W. J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J. F., Ring, B. Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S. L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Sempke, C. A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sincalir, B., Sperling, S., Stupka, E., Sugita, K., Sultana, R., Takekura, Y., Teki, K., Tammoja, K., Tan, S. L., Tang, S., Taylor, M. S., Tegner, J., Teichmann, S. A., Ueda, H. R., van Nimwegen, E., Verardo, R., Wei, C. L., Yagi, K., Yamashita, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S. M., Teasdale, R. D., Liu, E. T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J. S., Hume, D. A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Kono, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watanabe, A., Okamura-Oho, Y., Suzuki, H., Kawai, J., and Hayashizaki, Y.
PANTOM Consortium
The transcriptional landscape of the mammalian genome
Science 309 (5740), 1559-1563 (2005)
- JOURNAL
PUBMED
16141072
- REFERENCE
7
Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C. C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K. C., Hallinan, J., Mattick, J., Hume, D. A., Lipovich, L., Batalov, S., Engstrom, P. G., Mizuno, Y., Faghihi, M. A., Sandelin, A., Chalk, A. M., Mortazavi, T., Tabar, S., Liang, Z., Lenhard, B., and

OY	1279	AGCCTGCGCCGAGAGAAACCCCGCGGGTAAACAACGCGGGAGCGAGCCGCTCGGCG	1338
Db	1419	CTCCCTTGTGATTCGCGAGGCTCCCGTANTCAAGAGCACAAGCAGAGAGAGACGACGG	1478
OY	1339	CCCAAGAGAAAGAGAGAGGTGCTTAAAGATAGAACCCAGACAGATAAACTTCAGCCC	1398
Db	1479	ACCA--TATCATCTGACGCGCCCGACACAGTGGGCCATCTTCCCTGAGACCCCGTG	1536
OY	1399	CCGACGCAAGCAGAGCTCCCTGACTGGCACTCTCACTTTCCGACGTGTGTGAAC	1458
Db	1537	CTCGGGCTGGGTCCAGGCTTTTCTATGATATGTCTCAGCCCTTCCCTCAGGCTGAAG	1596
OY	1459	TCGACGGGGGAGTGGCCCTTGCCACAGTACCTTGTCTCCCACTGACCCCGACAC	1518
Db	1597	AGCCACTGAGCTGTGCGCTTGAGCTTGAGGCGCG-----CTGATGACAGCTGAG	1647
OY	1519	AAGCAGGCTCTCCCTTGATCAGCCCGCGCTGAATGACACAGGCTGCGCCCGAGCGAG	1578
Db	1648	CAGGTTGATCCCTGTGTCTCAGCCCTGTGGCTCAGTGGCGCTGTACAGCAGAGATGAC	1707
OY	1579	GATGAGATGAGGCGCGAGGAGAAGATTTCCCACTGATTAAGCGCTACTATAGTAA	1638
Db	1708	AACCAAGAAAGCAGAAACACAGACATGCCAGAAAGTGAAGCTTATTATAGCCAG	1767
OY	1639	GAAGTGTCTACCAACGAGCGAACAATATGTGAAGATCTGAAAGTTATCATCTTGTT	1698
Db	1768	GAGATTTCTGCTACAGAAACGAACTATCTGAAGGATTTAAGAAATATCATCTGTGGTT	1827
OY	1699	CAGAGCAGAGAGCAAGAGAGAGCGCCATGCCGAAAGCATGAAGAAAGTCTCATATCCG	1758
Db	1828	CGAGCCGTGCTGATCAGAGAGAGGCGCATGCTCGAGGCTTGATGGCCCTCTTTCTCC	1887
OY	1759	AATTTTGAACCTTTGCAAAATTTATCTAATTTTTCAGAGAAATTGACCAACGACTT	1818
Db	1888	AACATGATTCAGATCTACGATTTCCACAGAGGCTTCTTACAGAGTGAACAGAGCTG	1947
OY	1819	GCCCTGTGGGAAGCGCGCTCAAAATGCCA---AATCAGAGATTACAAAGAAATCGCGAT	1875
Db	1948	GCACTCTGGAAAGGCGCTCCAGATGCCACTTTAAAGATGATACACAGAAATCGGGAC	2007
OY	1876	GTCAATGCTGAAGAAACATTCAAGGCAATGAACACCTGGGGCTCACCTGTGAACACAGC	1935
Db	2008	ATCTCTCTCAGAGAAATGCTGTCAGTTAAAGAAATTTACTAGCTTCTTCAAGACACAT	2067
OY	1936	GAGGCTTGAAGGCGCTGAGAAATGGAATCAAGAGCTCCCGCGGCTGAGAACTTCTGC	1995
Db	2068	GAGTCTCTAACAGAACTGAAAAAGGCCACAAACACTGTAAAAAGCTGAGGAGCTTAC	2127
OY	1996	AGAGACTTTGAGCTGACAGAGGTGTATTACCTACCGCTCAACACCTTCTCTTGCGGCA	2055
Db	2128	AAAGATTTGAGCTCCAAAGGTGCTGTAACGTGCTTCAACACATTCCTGCTGAAGCC	2187
OY	2056	CTGCACCGGCTCATGCACTACAGAGAGTCTTGAAGCGGCTGTGCAAAACACACCCGCG	2115
Db	2188	GTCCAGAGGCTAGTCACTACCGTGTGTGCTGAGCCGCGTGTGTCTACCTACTCTCT	2247
OY	2116	AGCCACGCGCACTTCAGGAGCTGCGGACCGCGTTTGAGAGATCAACGAGATGTGGCA	2175
Db	2248	GAGCAGCGGACTATGCGGACTGCCATAGGCACTGAAGGCACTACAGAACTGACAC	2307
OY	2176	CAGCTCCACGGTACGATGATCAAGATGAGAAATTTCCAGAACTGCAAGACTCAAGAA	2235
Db	2308	GAGCTCCAGCAAAAGCTTAACCGGCTGAAAACTACAGAAATGAGCGAGCTACAGGA	2367
OY	2236	GATTGATTTGGCAATGCAATCTTGTGTTCCGGGAGGGAAGTTTATCTGCTGGCAGC	2295
Db	2368	GACCTGTGCGGTGTAAGAAACCTCAATCTCTGTGAGGAGATTAATCCGTGAAGGCTGC	2427
OY	2296	CTCAGCAAGCTCTCGGGAGAGGGGCTCAGAGAGCGCATGTTCTTCGTTCAAGAGCTC	2355
Db	2428	CTGCACAAAGCTTACCAAGAAAGGCTCTGACACAGAGAGATTTTTTTCTGTCTCAGATATG	2487
OY	2356	CTGCTATACAGAGCCGGGGCTGACGGCTTCATCATAGTTTAAAGTCAAGGACGCTC	2415

Db	2488	TTACTGTAATCAAGCAAAAGTGTACAGAGAGCCAGTCAATTTCCGGATCCGTGCTTCTT	2547
QY	2416	CCGGCTATATGCAATGACGATTTGAGAGAGCGAAGACGATGGGGGGTGGCCCACTGCTG	2475
Db	2548	CCACTCCGTGGCAATGCTGTGTAGAGAAAGTGAATGAAATGTCTGTTCTTTCATTTGCTTC	2607
QY	2476	ACCCTCGGGGCCACGGCGAGTCCATCATCTGTGGCCGCGACGTTCTCGGTCCGATGGAG	2535
Db	2608	ACCATCTATGACAGCTCAGAAAAAACAATTGTGGTACACACCAGCACTCCGGCTTAGAAAAAGAA	2667
QY	2536	AAGTGGGTGATGGACATCCAGATGGCCATTGACCTGGCGGAGAGAGACAGACGCCGCCG	2595
Db	2668	AAGTGGATGACGAGACCTGGAATGCAGCAATCCAGACACCAGACTATTCGGTGATCCACC	2727
QY	2596	CCTGAGTTCCTGGGCGACGAGCCGCCCTGACAAACAAGTTCCTCCGTGATGAAGCACCGCGCT	2655
Db	2728	CCAATGCTGCTGG---GAGGCCCGGTATATCTGTATACCTTAGATCTTCTGATGAAGTGC	2784
QY	2656	GACCAAGATCAGAGATGACCTTAGAGCCCTCCGCGCAATTCGCTGAGACGCGACAGCCCG	2715
Db	2785	TCTCTGGAAGAATCAGAAAGATGATGCAGAGAAACGGGGGCTCCCTGAGAGGGAAACAGCAG	2844
QY	2716	CACCGCGGCACACAAATGATGATCAGTGTGCTGGACACCGCAACACACACGCTTCCATGCTG	2775
Db	2845	CACCGGGCAATACAAACAAATGCAATGTGTGCTGATCCGTAAATACAAATGTGTCCAGACA	2904
QY	2776	GACTTACAGCATCGGAGTGGAGAAATCAGTGTGCTGAAACCTGCGTAGAGAAATTCAAAAAC	2835
Db	2905	GACCAACAGTGCAGCTGTTGGAGAACCAAGCTTTCAGAGATATCTGTGAGAAAGTTCAAGAC	2964
QY	2836	AGCAACGGGTGGCAAGAGCTGTGGTGGTGTTCACAAACTTGTGCTGTTCTTCTACAA	2895
Db	2965	AGTATAGGCTGGCAAGAAAGCTCTGGTGGTCTTTTACAACTTGTGCTTCTTATAAA	3024
QY	2896	TCAACACAGACAAATCATCCCTTGGCCAGCTGCTGCTCGCTACTGCTACCATC	2955
Db	3025	ACACATCAGAGATGACTACCCCTGCGCCAGCTCCCACTATGGGTTCACAGGTGAGGCTC	3084
QY	2956	CCCTCTGATGTCGAGAAACATCCAGAAAGACTAGCTGTTCAAGCTGCACCTCAAGTCCAC	3015
Db	3085	CCGAGGAGGCTGACAGCATTCACAAAGACTATGTCTTCAAGCTCCAAATTCAAATCTCAC	3144
QY	3016	GTTCTACTACTTTCAGAGGCGGAAAGCAGATACAGTTTCGAAAGGTGATGAGAGATGATCGC	3075
Db	3145	GTCTACTTCTTCCGGGGCTGAGAGCAATACAAATTTGAAAGGTGATGAGAGTATCAAA	3204
QY	3076	AGTCCACACAGCTTCTGCTGCGACCCCAAGTGTGAGCCACAAAGACTCTTT	3129
Db	3205	AGGGCCAGCAGCTCACACAGGAGAGACCCCAAGTTTCACTCAGGACTGCTCATCAT	3258
RESULT 6			
AK050860		3874 bp	mRNA
LOCUS			linear
DEFINITION		Mus musculus 9 days embryo whole body cDNA, RIKEN full-length	HTC 02-SEP-2005
ACCSSION		enriched library, clone: D030026M03 product: weakly similar to CDEP	
VERSION		[Homo sapiens], full insert sequence.	
KEYWORDS		AK050860.1 GI:26094164	
SOURCE		HTC; CAP trapper.	
ORGANISM		Mus musculus (house mouse)	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;	
JOURNAL		Sciuromiina; Muridae; Muridae; Murinae; Mus.	
PUBMED		Carninci, P. and Hayashizaki, Y.	
AUTHORS		High-efficiency full-length cDNA cloning	
		Meth. Enzymol. 303, 19-44 (1999)	
		10349636	
		2	
		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,	

TITLE	Itch,M., Komo,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
PUBMED	11042159
REFERENCE	
AUTHORS	3 Shibata,K., Itch,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitanai,T., Tashiro,H., Itch,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kaishiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The PANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6
AUTHORS	RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Core Team) and the FANTOM Consortium.
TITLE	Antisense transcription in the Mammalian Transcriptome
JOURNAL	Science 309, 1564-1566 (2005)
REFERENCE	7
AUTHORS	The PANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group).
TITLE	The Transcriptional Landscape of the Mammalian Genome
JOURNAL	Science 309, 1559-1563 (2005)
REFERENCE	8 (bases 1 to 3874)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Horii,P., Imocani,K., Ishii,Y., Itch,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takai-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
JOURNAL	Direct Submission
TITLE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsukuba-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:htp://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:htp://genome.gsc.riken.jp/ URL:htp://fantom.gsc.riken.jp/. Location/Qualifiers
FEATURES	1..3874
source	organism="Mus musculus" mol_type="mRNA" strain="C57BL/6J" db_xref="FANTOM_DB:D030026M03"

	Query Match	32.1%;	Score 1007.6;	DB 6;	Length 3874;	
	Beet Local Similarity	59.1%;	Pred. No. 5.7e-255;			
	Matches 1843;	Conservative 0;	Mismatches 1244;	Indels 29;	Gaps 6;	
	misc_feature					
	/note="putative					
	Weakly similar to CDEP (Homo sapiens) (SPT8.09x4F1,					
	evidence: FASTY, 53.7%ID, 99.7%length, match=3125"					
	1. .3874					
	/dev stage="9 days embryo"					
	/clone_id="Riken full-length enriched mouse cDNA library"					
	/cissue_type="whole body"					
	/db_xref="taxon:10090"					
	/clone="D030026M03"					
QY	CCGACCCCCAGGATTCAGCACTGGGGGCCCCGGAAAATTCGGGGATCACTACTTGGGAACGT	81				
DB	103 CCAACTTCAGGAGACACCTGGGGGGCCCAACCGCCATTTGAGTATACCTCGAGACCA	162				
QY	82 GGAAGAGAGCGCGCCCAACACCTTCAGAAAATCGCTGTCCATCAAAATCCAGATCTG	141				
DB	163 GAGCAGAGCGCTGTACCCAGAGATGACAGAGAAAGACATGGGCATCAAGTAGAGCTGCTG	222				
QY	142 GATGACACCCAGAGAGGCAATTGAAATTCACAAAGAGCTCCTGGAAAGGTCTGCTGGAT	201				
DB	223 GACAGCACTGATGATTTGACATGAGCCGAGGTGACCGGCGAGTCTTACT	282				
QY	202 GCAGTTTGCACACCACTTCACCTGTGGAAGTGATCATTTTGGCTCGAGATTTCTGAT	261				
DB	283 CAACTGTGGAAGCAATTGAAATTCATGAAATGTGACATCTTGGTCTGAGTTCAAGAT	342				
QY	262 CACAAAAAGATCACGGTGTGAGTCTCTTAAACCCATTGTGAACAGATTAGAAG	321				
DB	343 GTCCAGTCTCACTGGAATTTGGCTTGAACCTATGAAACCCATCATTTAGGCAAGTAGAAG	402				
QY	322 CCAAGAGACGCTGTGTTAGTTTGTGTGTAATTTCTTCCGCGTGCACACACAACACTC	381				
DB	403 CCAAAAAATGGGGTGCTTCGCTCGCCAGTAATAATTTTCCGCGTGCACCTGGTCACTG	462				
QY	382 CAAGAGAATTCACAAAGTACTGTGTGCGCTGCAGGTGAGAGAGACTTGGCTCAAGGC	441				
DB	463 CAAGAGAATTCACAAAGTACTGTGTGCGCTGCAGGTGAGAGAGACTTGGCTCAAGGC	522				
QY	442 AGGTTCACGTGATATGACACCAAGGCGAGCTCTCTGATTTCAACATTTGCAATCTGAG	501				
DB	523 CGCTGCATCTGCACAGGCGCAACCTGCAGCGCTTCTCAATCCACTCTTCGCACTGGAA	582				
QY	502 ATTGGGATTTTGTGTAAGCCTTGGACAGAGAGCACTTACAAAATAATTAATCATCT	561				
DB	583 ATCGGAATTTATGATGAAACCTCGAGATCGAAGAACCTCAAAAGCATGAAATCTGCC	642				
QY	562 CAGCAAGACGCACTAGAGACAAATATCTGGAATTTCACTAATACCATTAACATTGACAA	621				
DB	643 AACACAGAGAAATCTTAGAAAAGATCTAGACTTCATCAGAGGCAACAGCGGCAACT	702				
QY	622 CCACGAAATTCAGATTTCCAGCTCCTAGAGATTTGCCGTGGCTAGAGATGATGATC	681				
DB	703 CCGCAGAGTCAAGATTTCCAGGTGCTTGAATGTGCAGAGAAAGCTGGAAATGATATGCA	762				
QY	682 CGGTTGCAACCCGCGCAAGAGACAGGGAAGGACAGAGATCAATCTGCGCTTGCACAGC	741				
DB	763 AGTTTTCACATGGCTTCTGACAGAGGAAGGACCAAGATTAATCTAGCAATTTCTCATG	822				
QY	742 GGAATTTAGGTGTTTCAGGTTTTCATCAATATCAATGCTTTCAACTGGGCGCAAGGTCCG	801				
DB	823 GGTGTCTCGATGTTTCCAGGGGTACCAACCAATCAATCTTCAAATGCTTCCAAAGTCCGT	882				
QY	802 AAGCTGAGCTTCAAGAGAAAGCGTTTCTATCAAGCTCGGCGCAGATGCAATATGTGCG	861				
DB	883 AAATTAAGCTTCAAGAGAAAGGTTTCTTATCAAAATCCACCTGAGGTCCAGTGGGCC	942				
QY	862 TACCAAGATACCTTGAATTCGTGATGCGCATGTCGGGATTTCTGCAGATCCTTGTGAAA	921				

Dh 943 TACAGAGACAGTTAGATTCTTGTGGGTAGCAGAGATGAATGAAGAACTTCTGGAG 1002
Qy 922 ATCTGTGTGAACATCATGCTTCTTTAGACTTTTGAAGAGCCCAACCAAGCCAG 981
Db 1003 ATATGTGTGAATCACTACTTTTAACTCTCTACACAGCCTTAAGCCAAAGGCAAG 1062
Qy 982 CCGGTCTCTTTAGCCGGGGTCACTTCGTTCAAGTGTGGAGCTCAGAAAGCAAGTT 1041
Db 1063 GCTGTCTTCTTACAGCCAGGCTCTCTTCAAGATACGTGGAAGAACTCAGAAACAACTA 1122
Qy 1042 CTCGACTATGTAAAGAGAGACATAGAAGGTGAGTTTGAAGAAAGCAAGCAGAG 1101
Db 1123 GTAGATTATGTCAAAAGACGGTGAATGAAGAAATTCATACGAAGAACGGCAGATAG 1182
Qy 1102 ATTCT--ATTCTATCCGAGGCTTGTTCACAGCTTACAGAACTGAATTGGAAAGTCTG 1158
Db 1183 ACTCGGACATCTCTTCAATGCTCTGACTGTAGATGTGCTTAAACAGACGTCCTTACCC 1242
Qy 1159 GAGAGTCTCAGACAGACACAGCCTTAATTTGAGAGAGGTGCGGAATCTCAGAGGGAGC 1218
Db 1243 GATGGCTTGAAGACTTGTGCTCTCTGTCTTCAAGAAATGTCTCTTTTATCCAC----- 1297
Qy 1219 CAGAGCTGCGCGAGAGAAAGAACCGAAGGTTTCGCGCGGAGACCGGGGTGCGACCG 1278
Db 1298 -----CCCTTAGTTCTTCTTGTCTCTGCTGCTGCCCAATTGAAAGGACAGCAGAG 1352
Qy 1279 AGCCCTGCGCGAGAGAGAGAGCCCGCGGGTAAACAAGAGCGAGCGAGCCGCTCGGCG 1338
Db 1353 CTCCTCTGTGATTCGCGAGGCTCCCTCATCAAGAGCAGCAGCAGAGAGAGCAGCAGG 1412
Qy 1339 CCCACGAGAGAGAGAGAGAGGTGTTAAGGATAGAGCCACAGAGATTAACCTCAGCCC 1398
Db 1413 ACCA--TCATCATCTGAGCGGCCCAACACAGTGGGCCCATCTCCCTGAGCCCCCGTG 1470
Qy 1399 CCGAGCCGAAGCAGAGGCTCCCTGACTGAGCTGCTCACTTTCGAGCTGTCTGTAGAC 1458
Db 1471 CTCGCGCTGTGCTCAGGCTTTTCTATGAGATGCTCTCAGCCTTCTCCCTCAGCCTGAG 1530
Qy 1459 TCGAGAGGGAGAGTGGCCCTGCGCAAGTGAACCTTGTCTCCAACTGAGCCCGACACC 1518
Db 1531 AGCCACTGAGGCTGTGCTCTGAGCTTCAAGCCCG-----CTGAGTACAGCTGAG 1581
Qy 1519 AAGAGGCTCTCCCTTATGAGCCGCTGTGAATGACAGAGCCTGCCCGGAGCGAGC 1578
Db 1582 CAGGTGCACTCCCTGTGCTGAGCCCTGTGCTCAGTGTGCTGAGCGCTGTACAGCCAGATGAC 1641
Qy 1579 GATGAGATGAGGCGCCGAGAGAGATTCCTCACTGAATGAAGCTTATATAGTAAAG 1638
Db 1642 AACCAAGAGAGAGAGAAACAACACATGCAAGATAGAGCTATTTTATATAGCAG 1701
Qy 1639 GAATGTCTAACACCGAGCGAATATCTGAGAGATCTCGAAGTTATCATCTGTGTTT 1698
Db 1702 GAGATTCTCGTACAGAACGAACTATCTGAGAGATTAGAAAGTTATCATCTGTGTTT 1761
Qy 1699 CAGAGCAGATGAGCAAGAGAGAGAGCGCATGCGGAGACCTGAAGAAATCTCATATTCG 1758
Db 1762 CGAGGCTGTGATCAAGAGAGAGAGAGCGATGCTGACGCCCTGATGAGCCCTGCTTTTCC 1821
Qy 1759 AATTTGAACCTTTCACAAATTTTCATATTAATTTTTCAGAGAAATTTGAGCAAGCTT 1818
Db 1822 AACATTATCAAGTCTACAGATTCCACAGAGGCTTCTTCAAGAGGTGAACAGAGCTG 1881
Qy 1819 GCCCTGTGGAAGGCGCTCAATATGCCA--AATCAGAGATTACCAAGATTTGGGAT 1875
Db 1882 GCACTCTGGGAAGGCGCTCAGTGCACCTTAAAGGTATCAACAGCATATGGGAGC 1941
Qy 1876 GTATGCTGAAGACATTTACAGGCGATGAGCACTGGGCGCTCCTGTGGAAGCAGAGC 1935
Db 1942 ATCTCTCTCAGAGACATGCTGCAATTAAGAAATTTTACTAGCTACTTCAAGACAGAT 2001
Qy 1936 GAGGCTCTGAGGCGCTGTGAGAAATGAATCAAGAGCTCCCGCGCTGTGAGAACTTCTGC 1995
Db 2002 GAGGTCTTAACAGAACTGGAAAAGGCCAABAAACATGTAAAAAAGCTGTGAGGCAGTCTAC 2061

Qy 1996 AGAGCTTTGAGCTGAGAGAGGTGTGTACTACCGCTCAACACTTCTCTGCGGCGCA 2055
Db 2062 AAAGATTGTAGCTCAAAAGGTGTGTGTACTGTCTCAACAACTTCTGCTGAAGCC 2121
Qy 2056 CTGACCGGCTCATGACATCAAGACAGGTCTTGAAGCGGCTGTGCAACACACCGCGC 2115
Db 2122 GTCCAGAGGCTATCCATACCGTCTGTGCTGAGCGGCTGTGTGTGCTCATCTGTCT 2181
Qy 2116 AGCCAGCGGACTTCAGGAGCTGCCAGCGCTTTTGGCAGAGATCACAGAGATGTGGCA 2175
Db 2182 GGGCACCGGCATATGCTGCGACTGCAAGAGGCACTGAAGGCACTGACAGAAAGTACACACC 2241
Qy 2176 CAGTCCACGCTACATGATCAAGATGAGAGAAATTTCCAGAAAGCTCAGAA 2235
Db 2242 GAGCTCAGAAAGCTTACCGGCTGGAAGAACTTACAGAAATTTGACGAGTCAAGCGA 2301
Qy 2236 GATTTGATTGGATTGCAATCTTGTGTGTTCCGGGAAGGAGTTCACTCGTGGGAGC 2295
Db 2302 GACTGGTGGTGTAGAAACCTCATTTGCTCTGGAGGAGTTTATCGGTGAGGCGTGC 2361
Qy 2296 CTCAGAGCTCTCGGGGAAGGGGCTTCAGAGCGCATGTTCTTCTGTTCAAAGACGTC 2355
Db 2362 CTCACAGCTCAACAAAGAGGCGCTGCAACAGAGATGTTTTTCTGTCTCAGATATG 2421
Qy 2356 CTGCTATACAGAGCGCGGCGCTGAGCGCTCCAATCAAGTTTAAAGTCCAGGGGAGCTC 2415
Db 2422 TTGCTGTATCAAGCAAAAGTGTCAAGAGCCAGATATTTCGGAATCGTGGCTTCTT 2481
Qy 2416 CCGCTCATGAGTGAAGATTTGAGAGAGCGAAGACAGATGGGGGGTCCCACTGCTG 2475
Db 2482 CCACTCGGTGGCATGTCTGTATGAAGAAAGTGAAGATGAATGTGTCTCTCATTTGCTTC 2541
Qy 2476 ACCCTCGGGGCGAGCGGCGATCATCTGTGCGCGCAGTTCTGTGCTCGAGATGAG 2535
Db 2542 ACCATCATGCACTCAGAAACAAATGTGTGTAGCAGCCAGCACTCGGCTAGAAAGGAA 2601
Qy 2536 AATGTGTTAGAGACATTCAGATGCGCATTTGACCTGGCGGAGAAAGACAGACCCCGCC 2595
Db 2602 AAGTGAATGAGAGACCTGAATGAGACATCAAGCAGCAAGCTATCGGTACTACCC 2661
Qy 2596 CTTGAGTCTCTGCGCAGACAGCCCGCTGACAAAGTCCCTGATGAAGCAGCGGCGT 2655
Db 2662 CAGTGTGTGGAGAGGCCCGGTGTATCTGTATCCCTGATCTTGTATGAAGTCTCT 2721
Qy 2656 GACCAAGATCAGAGATGACTGAGCGCTGCGGACATGCTGTGAGCGCCAGGCCCG 2715
Db 2722 CTGAAAGATCAGAAAGATGTGAGGAAAGCCGGGCGTGCCTGTGAGGGGAAACAGCCAG 2781
Qy 2716 CACCGCGG--CAACAAATGTGTGCTGTGCGGACACCGCAACCAAGGCTCTCAATG 2773
Db 2782 CACCGGGGCCCAATCAACATGATGTGTGTGTGTACGTTAATCAAGGTGTCCAGAG 2841
Qy 2774 TGAATTCAGCATCGAGTGAAGATCAGTTGTCTGGAACCTGTGAGGAATTCAAAA 2833
Db 2842 CAGATCAGATGAGCTGTGTTGAACCAAGCTTTCAAGATATCTGTGTAAGAAATTCAGA 2901
Qy 2834 ACAGCAACGGGTGCAAGAGCTGTGGGTGTTCACAAACTTCTGCTGTCTTCTTACA 2893
Db 2902 ACAGTAATGCTGGCAGAACTCTGGGTGTCTTTACAACTTCTGCTGTCTTCTATATA 2961
Qy 2894 AATCAACAGAGCAATCATCCCTTTCAGAGCGCTGCGTCTGCTCTGCTATCTGCTACCA 2953
Db 2962 AAACACATCAGATGACTACCCCTGTGCAAGCTCCCATATGTGGTTAAGAGGTGAGCC 3021
Qy 2954 TCCCTCTGAGTCCAGAAACATCCAGAAAGACTACAGTTCAGAGTTCAGTCAAGTCCC 3013
Db 3022 TCCCGAGAGGCTGACAGATCAACAAAGCTATGTCTTCAAGCTCAATTCAAATTCG 3081
Qy 3014 ACGTCTATCTTCAAGGCGGAAAGAGATGACAGTTCCAAAGGTGATGAAGTATCC 3073
Db 3082 ACGTCTATCTTCTCGGGCTGAGAGCAAGTACATTTGAAGGTGATGAGAGTCAATCA 3141

QY 3074 GCAGTCCACCAAGCTGTGCTCGGACCCCAAGCTGTGAGCCACAAAGTCTCTT 3129
 DB 3142 AAGGGCCAGCACTCAACAGGAGACCCCAAGTTTCACTCAAGACTGTCTACAT 3197
 RESULT 7
 LOCUS BX364606 968 bp mRNA linear EST 08-APR-2004
 DEFINITION BX364606 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 CDNA clone CS0DC027YA10 5-PRIME, mRNA sequence.
 ACCESSION BX364606 GI:46289717
 VERSION BX364606.2
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 968)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 5, 2003 this sequence version replaced gi:30370819.
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7229.f

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdnas=CS1AC007Z8060P1&c=7229.f.
 Location/Qualifiers

FEATURES
 source
 1..968
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC027YA10"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
 ORIGIN

Query Match 30.5%; Score 957; DB 4; Length 968;

Best Local Similarity 99.9%; Pred. No. 9,9e-242;

Matches 968; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1062 AGGACATTAAGAGGTGAGTTTGAAGAAAGCAACAGAGATTCTTATCCGAGCCT 1121
 DB 1 AGGACATTAAGAGGTGAGTTTGAAGAAAGCAACAGAGATTCTTATCCGAGCCT 60
 QY 1122 TGGTTACAGCCTTACGAAGTCTGAGAGTCTGAGAGCTTCTCAGAGAGCCAG 1181
 DB 61 TGGTTACAGCCTTACGAAGTCTGAGAGTCTGAGAGCTTCTCAGAGAGCCAG 120
 QY 1182 CCTTACATTTGAGAGAGTGCAGATCTCCAGGGGGCCAGAGTGCCTGGGAGAAAGN 1241
 DB 121 CCTTACATTTGAGAGAGTGCAGATCTCCAGGGGGCCAGAGTGCCTGGGAGAAAGN 180
 QY 1242 ACCGAAAGTTTCCGCGGGGAGCCGGGGGTGCACCCGAGCCTTGGCCGAGAGAAAGCC 1301
 DB 181 ACCGAA-GTTTCCGCGGGGAGCCGGGGGTGCACCCGAGCCTTGGCCGAGAGAAAGCC 239
 QY 1302 CGCGGGTAAACAAGCGGAGCGAGCCGCTCGGGCCCAAGAGAGAAAGAGAGGT 1361
 DB 240 CGCGGGTAAACAAGCGGAGCGAGCCGCTCGGGCCCAAGAGAGAGAGAGGT 299

QY 1362 CGTTAAGATTAAGACCCAGACAGATTAACCTTCAGGCCCCCGACGACAGAGCTCCCT 1421
 DB 300 CGTTAAGATTAAGACCCAGACAGATTAACCTTCAGGCCCCCGACGACAGAGCTCCCT 359
 QY 1422 GACTGGCAGTCTCTACCTTTCCGAGCTGTCTGTAACTTGGCAGGGGGAGGTGCCCTTGC 1481
 DB 360 GACTGGCAGTCTCTACCTTTCCGAGCTGTCTGTAACTTGGCAGGGGGAGGTGCCCTTGC 419
 QY 1482 CAACGTGACCTTGTCTCCCACTGAGCCCGACACCAAGAGAGGCTCTCCCTTATCAG 1541
 DB 420 CAACGTGACCTTGTCTCCCACTGAGCCCGACACCAAGAGAGGCTCTCCCTTATCAG 479
 QY 1542 CCCGCTGTGAATGAACAGAGCCTGCCCCCGAGCGAGCAATGAGATGAGGGCCGAGAA 1601
 DB 480 CCCGCTGTGAATGAACAGAGCCTGCCCCCGAGCGAGCAATGAGATGAGGGCCGAGAA 539
 QY 1602 GAGATTTCCCACTGATTAAGGCTACTTCACTAAGAAAGTGTCTACCAACGAGAGAAC 1661
 DB 540 GAGATTTCCCACTGATTAAGGCTACTTCACTAAGAAAGTGTCTACCAACGAGAGAAC 599
 QY 1662 ATATCTGAAGGATCTCGAAGTTATCATCTGATGTTTCAGACACAGTGAAGAGAGA 1721
 DB 600 ATATCTGAAGGATCTCGAAGTTATCATCTGATGTTTCAGACACAGTGAAGAGAGA 659
 QY 1722 CGCATGCGGAGAGCACTGAAAAGTCTCATATTCGGAATTTTGAACCTTTGCACAAAT 1781
 DB 660 CGCATGCGGAGAGCACTGAAAAGTCTCATATTCGGAATTTTGAACCTTTGCACAAAT 719
 QY 1782 TCACTAATTTTTCACAGAAATTGAGCAAGACTTGCCTGTGGAAAGGCGCTCAAA 1841
 DB 720 TCACTAATTTTTCACAGAAATTGAGCAAGACTTGCCTGTGGAAAGGCGCTCAAA 779
 QY 1842 TCGCCAAATCAGAGATTACCAAGAATCGGCGATGCTGTAAGAAATTCAGAGGACAT 1901
 DB 780 TCGCCAAATCAGAGATTACCAAGAATCGGCGATGCTGTAAGAAATTCAGAGGACAT 839
 QY 1902 GAGGACCTGGGGGCTGACCTGTGAAAGCAGAGAGCCTTGAAGGCCCTTGAAGATG 1961
 DB 840 GAGGACCTGGGGGCTGACCTGTGAAAGCAGAGAGCCTTGAAGGCCCTTGAAGATG 899
 QY 1962 AATCAAGAGCTCCCGGCGCTGTGAGAACTTCTGCAAGACTTTTGAAGCTTCAAGAGTGTG 2021
 DB 900 AATCAAGAGCTCCCGGCGCTGTGAGAACTTCTGCAAGACTTTTGAAGCTTCAAGAGTGTG 959
 QY 2022 TTACCTACC 2030
 DB 960 TTACCTACC 968

RESULT 8
 BX354921 1000 bp mRNA linear EST 23-APR-2004
 LOCUS BX354921
 DEFINITION BX354921 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 CDNA clone CS0DC027YA10 5-PRIME, mRNA sequence.
 ACCESSION BX354921
 VERSION BX354921.2 GI:46553760
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE 1 (bases 1 to 1000)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 5, 2003 this sequence version replaced gi:30383921.
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7229.f

For more information about this cluster, see <http://www.genoscope.cns.fr/cdnas=CS0PC027BA050P1&c=7229.f>.

FEATURES

source

```

1. 1000
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="CS0DC027YA10"
   /issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
   /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
   /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

```

Query Match      30.0%; Score 940.2; DB 4; Length 1000;
Best Local Similarity 99.2%; Pred. No. 2.9e-237;
Matches 961; Conservative 4; Mismatches 2; Indels 2; Gaps 2;

QY 1062 AGGACATAGAGAGGTGCGATTGTAAGAGAGACAGACAGAAATTCATTCTATCCGAGCCCT 1121
DB 1 AGGACATAGAGAGGTGCGATTGTAAGAGAGACAGACAGAAATTCATTCTATCCGAGCCCT 60
QY 1122 TGGCTTCACAGCCTTACAGAACTGAATTCGGAAGTGTCTGAGAGAGTCTTCAGACAGACACAG 1181
DB 61 TGGCTTCACAGCCTTACAGAACTGAATTCGGAAGTGTCTGAGAGAGTCTTCAGACAGACACAG 120
QY 1182 CCTTACCTTTGGAGAGAGTGTCCGAAGTGTCCGAGGGGGCCAGAGTGTCCGGCGAGAGAGAGGA 1241
DB 121 CCTTACCTTTGGAGAGAGTGTCCGAAGTGTCCGAGGGGGCCAGAGTGTCCGGCGAGAGAGAGGA 180
QY 1242 ACCGAAGTGTTCGGCGGGAGCGCGGGAGTGCACCCGAGCCCTGCGCGGAGAGAGAGAGCC 1301
DB 181 ACCGAAGTGTTCGGCGGGAGCGCGGGAGTGCACCCGAGCCCTGCGCGGAGAGAGAGAGCC 239
QY 1302 CGCGGGTAAAGAGAGAGCGAGCGAGCGCGCTCGCGCCGCCACGAGAGAGAGAGAGAGT 1361
DB 240 CGCGGGTAAAGAGAGAGCGAGCGAGCGCGCTCGCGCCGCCACGAGAGAGAGAGAGAGT 299
QY 1362 CGTTAAGAGATAGAGACCAAGCAGAGATTAACCTCAGCCCCCGGAGAGAGAGAGAGTCTCCT 1421
DB 300 CGTTAAGAGATAGAGACCAAGCAGAGATTAACCTCAGCCCCCGGAGAGAGAGAGAGTCTCCT 359
QY 1422 GACTGGCAGTCTCACTTCCGAGCTGTCTGTGAATCTCGAGGGGGAGAGTGGCCCTGCG 1481
DB 360 GACTGGCAGTCTCACTTCCGAGCTGTCTGTGAATCTCGAGGGGGAGAGTGGCCCTGCG 419
QY 1482 CAAAGTACCTTGTCTCCCAACTGAGCCCGGACACCAAGAGAGGCTCTCCCTTGATCAG 1541
DB 420 CAAAGTACCTTGTCTCCCAACTGAGCCCGGACACCAAGAGAGGCTCTCCCTTGATCAG 479
QY 1542 CCGCGTGTCTAATAGCAGAGCGCTGCCCCCGGACGAGAGATGAGAGGCGCGAGAGAA 1601
DB 480 CCGCGTGTCTAATAGCAGAGCGCTGCCCCCGGACGAGAGATGAGAGGCGCGAGAGAA 539
QY 1602 GAGATTCCTCAAGTATTAAGGCTTCAATAGCTTAAGAGAGTGTCTACACCGAGCGAGAC 1661
DB 540 GAGATTCCTCAAGTATTAAGGCTTCAATAGCTTAAGAGAGTGTCTACACCGAGCGAGAC 599
QY 1662 ATATCTGAAGAGATCTCGAAGTTATCACTTGTGTGTTTCAAGACACAGTGAAGAGAGGA 1721
DB 600 ATATCTGAAGAGATCTCGAAGTTATCACTTGTGTGTTTCAAGACACAGTGAAGAGAGGA 659
QY 1722 CGCGATGCGCGGAAGACATGAAAGTCTCATATTTCCGAATTTTGAACCTTTGCAAAAT 1781
DB 660 CGCGATGCGCGGAAGACATGAAAGTCTCATATTTCCGAATTTTGAACCTTTGCAAAAT 719
QY 1782 TCATATTAATTTTCTCAAGAAATTAAGACAGACTTGCCCTGTGGAGAGGCGGCTCAAA 1841

```

```

DB 720 TCATATTAATTTTCTCAAGAAATTAAGACAGACTTGCCCTGTGGAGAGGCGGCTCAAA 779
QY 1842 TGGCCAAATAGAGATTAACAAAGAAATCGCGATGTATCTGTAAGAACTTCAGGGCAT 1901
DB 780 TGGCCAAATAGAGATTAACAAAGAAATCGCGATGTATCTGTAAGAACTTCAGGGCAT 839
QY 1902 GAAGCAGCTGCGGGCTCACTGTGGAAGACAGAGAGGCC-TTGAAGCCCTGGAGAAATG 1960
DB 840 GAAGCAGCTGCGGGCTCACTGTGGAAGACAGAGAGGCCCTTTGAAGCCCTGGAGAAATG 899
QY 1961 GAATCAAGAGCTCCGCGCGCTGAGAGAACTTTCAGAGAGACTTGAAGCTGAGAGAGTGT 2020
DB 900 GAATCAAGAGCTCCGCGCGCTGAGAGAACTTTCAGAGAGACTTTCAGAGAGCTTCAGAAATGT 959
QY 2021 GTTACTTAC 2029
DB 960 GTTACTTAC 968

```

```

RESULT 9
B0881649
LOCUS
DEFINITION
ACCESSION
VERSION
B0881649.1 GI:22273657
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 912)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM2533 row: n column: 12
High quality sequence stop: 684.
Location/Qualifiers
1. 912
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:6339275"
   /issue_type="neuroblastoma, cell line"
   /lab_host="DH10B (phage-resistant)"
   /clone_lib="NIH_MGC_47"
   /note="Organ: Brain; Vector: pOTB7; Site_1: XhoI; Site_2:
   EcoRI; cDNA made by oligo-dT priming. Directionally
   cloned into EcoRI/XhoI sites using the following 5'
   adaptor: GGACGAG(G). Size-selected >500bp for average
   insert size 1.8kb. Library constructed by Ling Hong in
   the laboratory of Gerald M. Rubin (University of
   California, Berkeley) using ZAP-cDNA synthesis kit
   (Stratagene) and Superscript II RT (Life Technologies).
   Note: this is a NIH_MGC Library."

```

FEATURES

source

```

1. 912
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:6339275"
   /issue_type="neuroblastoma, cell line"
   /lab_host="DH10B (phage-resistant)"
   /clone_lib="NIH_MGC_47"
   /note="Organ: Brain; Vector: pOTB7; Site_1: XhoI; Site_2:
   EcoRI; cDNA made by oligo-dT priming. Directionally
   cloned into EcoRI/XhoI sites using the following 5'
   adaptor: GGACGAG(G). Size-selected >500bp for average
   insert size 1.8kb. Library constructed by Ling Hong in
   the laboratory of Gerald M. Rubin (University of
   California, Berkeley) using ZAP-cDNA synthesis kit
   (Stratagene) and Superscript II RT (Life Technologies).
   Note: this is a NIH_MGC Library."

```

ORIGIN

```

Query Match      28.2%; Score 883.4; DB 3; Length 912;
Best Local Similarity 98.7%; Pred. No. 3.2e-222;
Matches 901; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 544 AAAATTAATATACCTCTCAAGAGAGCACTAGAGACAAATCGTGAATTTACCAT 603

```

```
Db      1 AAAAAAAAACTACTACTGCAAGACGCACTAGAGACAAATCGGAAATTTCAACAT 60
Qy      604 AACCACTTGGACAAACACCAAGCAATGATTTCCAGTCTCTAAGATATGCCCTCGG 663
Db      61 AACCAATTTGGACAAACACCAAGCAATGATTTCCAGTCTCTAAGATATGCCCTCGG 120
Qy      664 CTAGAGATGATGAAATCCGGTTGCAACCCGGCCAAAGACAGGAGGACGAAGATCAAT 723
Db      121 CTAGAGATGATGAAATCCGGTTGCAACCCGGCCAAAGACAGGAGGACGAAGATCAAT 180
Qy      724 CTGGCCGTGGCCAAACCGGAAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCCCTTC 783
Db      181 CTGGCCGTGGCCAAACCGGAAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCCCTTC 240
Qy      784 AACTGGGCGCAAGGTGGGGAAGCTGAGCTTCAAGAGAAAGGCTTTCTCAATCAAGCTCCGG 843
Db      241 AACTGGGCGCAAGGTGGGGAAGCTGAGCTTCAAGAGAAAGGCTTTCTCAATCAAGCTCCGG 300
Qy      844 CCAGATGCCAATAGTGGTACCAAGATACCTTGAATTCCTGATGGCCAGTCCGGATTTTC 903
Db      301 CCAGATGCCAATAGTGGTACCAAGATACCTTGAATTCCTGATGGCCAGTCCGGATTTTC 360
Qy      904 TGCAGTCTCTTGTGAAAATCTGTGTGAACATCATGCTCTTTTGAATTTTGAAGAG 963
Db      361 TGCAGTCTCTTGTGAAAATCTGTGTGAACATCATGCTCTTTTGAATTTTGAAGAG 420
Qy      964 CCGAAACCAAAAGCCCAAGCCCGTCTTTAGCCGGGGGTCATCATTTGGTTCAGTGTGT 1023
Db      421 CCGAAACCAAAAGCCCAAGCCCGTCTTTAGCCGGGGGTCATCATTTGGTTCAGTGTGT 480
Qy      1024 CGGACTCAGAGAGAGGTTCTGACTATGTTAAAGAAAGAGACATAGAAGGTGCAGTTT 1083
Db      481 CGGACTCAGAGAGAGGTTCTGACTATGTTAAAGAAAGAGACATAGAAGGTGCAGTTT 540
Qy      1084 GAAAGAGACAGCAAGATTCATTCTATCCGAGGCTTGTCAAGCTTACAGAACTG 1143
Db      541 GAAAGAGACAGCAAGATTCATTCTATCCGAGGCTTGTCAAGCTTACAGAACTG 600
Qy      1144 AATTCGAGAGTGTGAGAGAGTCTCAAGAGACCAAGCCTTTCATTGGAGAAAGTGGC 1203
Db      601 AATTCGAGAGTGTGAGAGAGTCTCAAGAGACCAAGCCTTTCATTGGAGAAAGTGGC 660
Qy      1204 GAATCTCCAGGGGGGCGAGAGCTCCGGCGAGAGAAAGAACGAAGTTTCCGCGGGAG 1263
Db      661 GAATCTCCAGGGGGGCGAGAGCTCCGGCGAGAGAAAGAACGAAGTTTCCGCGGGAG 720
Qy      1264 CCGGGGTCCGACCCGAGCCTCTGCGCGAGAGAAAGCCCGGGGTAAACAAGCAGCGGAG 1323
Db      721 CCGGGGTCCGACCCGAGCCTCTGCGCGAGAGAAAGCCCGGGGTAAACAAGCAGCGGAG 780
Qy      1324 GAGAGCCGCTCCGGCGCCCAAGAGAGAAAGAGAGAGTGTGTAAGATAGAGACCAAGCAG 1383
Db      781 GAGAGCCGCTCTTCCGCGCCAGAGAGAGAAAGAGAGTGTGTAAGATAGAGACCAAGCAG 840
Qy      1384 AGTAAACCTTCAAGCCCGCGAGCAAGCAAGGCTCCCTGACTGGAGTCTTCACTTTTC 1443
Db      841 AGTAAACCTTCAAGCCCGCGAGCAAGCAAGGCTCCCTGACTGGAGTCTTCACTTTTC 899
Qy      1444 GAGCTGTCTGTGA 1456
Db      900 GAGCTGTCTGTGA 912

RESULT 10
BUB31960          904 bp      mRNA      linear      EST 18-OCT-2002
LOCUS             BUB31960
DEFINITION       AGENCOURT 10485886 NIH MGC 109 Homo sapiens cDNA clone
IMAGE:6672254 5', mRNA sequence.
ACCESSION       BUB31960.1  GI:24120779
VERSION          BUB31960.1
KEYWORDS         EST.
SOURCE           Homo sapiens (human)
```

```
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo
REFERENCE      1 (bases 1 to 904)
AUTHORS       NIH-MGC http://mgc.nci.nih.gov/.
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cga@bbs-rcmail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: LNCM2952  row: h  column: 14
              High quality sequence stop: 721.
              Location/Qualifiers
                1..904
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6672254"
                /tissue_type="teratocarcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_1ib="NIH MGC 109"
                /note="Organ: ovary; Vector: pORF7; Site_1: EcoRI; Site_2:
                XhoI; cDNA made by oligo-dT priming. Directionally cloned
                into EcoRI/XhoI sites using the following 5' adaptor:
                GGCAAGAG(G). Library constructed by Ling Hong in the
                laboratory of Gerald M. Rubin (University of California,
                Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                Superscript II RT (Life Technologies). Note: this is a
                NIH-MGC library."

ORIGIN
Query Match      27.9%; Score 873.2; DB 3; Length 904;
Best Local Similarity 98.6%; Pred. No. 1.6e-219;
Matches 891; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

Qy      967 AAACCAAGCCCAAGCCGCTCTTTAGCCGGGGTCAATCTTGGTTCAGTGTGCG 1026
Db      1 AAACCAAGCCCAAGCCGCTCTTTAGCCGGGGTCAATCTTGGTTCAGTGTGCG 60
Qy      1027 ACTCAGAGACAGGTTCTGACTATGTTAAAGAGAGACATAGAAGGTGCAGTTTGA 1086
Db      61 ACTCAGAGACAGGTTCTGACTATGTTAAAGAGAGACATAGAAGGTGCAGTTTGA 120
Qy      1087 AGAAGACAGAGAGATTCATTCTATCCGAGCCTTGTCTTCAAGCTTACAGAACTGAAT 1146
Db      121 AGAAGACAGAGAGATTCATTCTATCCGAGCCTTGTCTTCAAGCTTACAGAACTGAAT 180
Qy      1147 TCGGAAAGTCTGAGAGAGTCTCAAGAGACCAAGCCTTTCATTGGAGAGGTGCCGA 1206
Db      181 TCGGAAAGTCTGAGAGAGTCTCAAGAGACCAAGCCTTTCATTGGAGAGGTGCCGA 240
Qy      1207 TCTCCAGGGGGGCGAGAGCTCCGGCGAGAGAAAGCAAGGTTTCCGCGGGAGCCG 1266
Db      241 TCTCCAGGGGGGCGAGAGCTCCGGCGAGAGAAAGCAAGGTTTCCGCGGGAGCCG 300
Qy      1267 GGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTAAACAAGCAGCGAGCA 1326
Db      301 GGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTAAACAAGCAGCGAGCA 360
Qy      1327 GCCGCTTCGGCCGCCCAAGAGAGAAAGAGAGTGTGTTAAGATAGAGACCAAGCAGAGT 1386
Db      361 GCCGCTTCGGCCGCCCAAGAGAGAAAGAGAGTGTGTTAAGATAGAGACCAAGCAGAGT 420
Qy      1387 AAACCTCAAGCCCGCGAGCAAGCAAGGCTCCCTGACTGGAGTCTTCACTTTCCAG 1446
Db      421 AAACCTCAAGCCCGCGAGCAAGCAAGGCTCCCTGACTGGAGTCTTCACTTTCCAG 480
```


1447 CTGCTGTAAGTGGAGGAGTGGCCCTGCGAAGCTGCTGCTCCCAACCTG 1506
Db CTGCTGTAAGTGGAGGAGTGGCCCTGCGAAGCTGCTGCTCCCAACCTG 540
Qy 1507 AGCCCCGACACCAAGCAGGCTCTCCCTTGATGACGCCGCTGCTGAATGACGAGGCTGC 1566
Db 541 AGCCCCGACACCAAGCAGGCTCTCCCTTGATGACGCCGCTGCTGAATGACGAGGCTGC 600
Qy 1567 CCCCCGACGACATGAGGATGAGGCGCGAGAAAGATTTCCCACTGATTAAGCGTAC 1626
Db 601 CCCCCGACGACATGAGGATGAGGCGCGAGAAAGATTTCCCACTGATTAAGCGTAC 660
Qy 1627 TTCTAGCTAGAGAGTGTCTACACGAGGGAACATCTGAGAGATCTGAGATTATC 1686
Db 661 TTCTAGCTAGAGAGTGTCTACACGAGGGAACATCTGAGAGATCTGAGATTATC 720
Qy 1687 ACTTCGTGTTTCAGAGCAGTGAAGCAAGAGAGAGCGCATGCGGAAGCACTGAAAAGT 1746
Db 721 ACTTCGTGTTTCAGAGCAGTGAAGCAAGAGAGAGCGCATGCGGAAGCACTGAAAAGT 780
Qy 1747 CTCTATATCCCGAATTTTGAACCTTTGCAAAATTTCTACTAATTTTCTCAGGAAAT 1806
Db 781 CTCTATATCCCGAATTTTGAACCTTTGCAAAATTTCTACTAATTTTCTCAGGAAAT 840
Qy 1807 GAGCAAGCACTTGCC---TGTGGAAAGCGCGCTCAATGCCAAATCAGAGATTACCA 1863
Db 841 GAGCAAGCACTTGCCCTGCTGAGGAGAGCGCGCTCAATGCCAAATCAGAGATTACCA 900
Qy 1864 AGAA 1867
Db 901 NGAA 904

RESULT 11
BM551766 1026 bp mRNA linear EST 20-FEB-2002
LOCUS BM551766
DEFINITION AGENCOURT_6575294 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5479350
5', mRNA sequence.
ACCESSION BM551766
VERSION BM551766.1 GI:18789071
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1026)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-ismail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2000 row: h column: 07
High quality sequence stop: 675.
Location/Qualifiers

FEATURES

Source
1..1026
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5479350"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_98"
/note="Organ: brain; Vector: POT87; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'
adapter: GGACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

Query Match 27.6%; Score 863.8; DB 2; Length 1026;
Best Local Similarity 96.7%; Pred. No. 5.1e-217;
Matches 916; Conservative 0; Mismatches 22; Indels 9; Gaps 3;
ORIGIN
Qy 1997 GAGACTTGGAGCTCAGAGAGTGTGTACCTACCGCTCAACCTTCTCTGCGGCAC 2056
Db 1 GAGACTTGGAGCTCAGAGAGTGTGTGTACCTACCGCTCAACCTTCTCTGCGGCAC 60
Qy 2057 TGACCGGCTCATATGCACTACAGCAGGCTCTGAGCGGCTGTGCMAACACACCGCGCA 2116
Db 61 TGACCGGCTCATATGCACTACAGCAGGCTCTGAGCGGCTGTGCMAACACACCGCGCA 120
Qy 2117 GCCACGCCGACTTCAAGGAGTCTCCGAGCGCTTTGGCAGAGATCAAGAGATGTGGCAC 2176
Db 121 GCCACGCCGACTTCAAGGAGTCTCCGAGCGCTTTGGCAGAGATCAAGAGATGTGGCAC 180
Qy 2177 AGCTCCAGGTTACGATGATCAAGATGAGAAATTTCCAGAGCTGCAAGAACTCAAGAAAG 2236
Db 181 AGCTCCAGGTTACGATGATCAAGATGAGAAATTTCCAGAGCTGCAAGAACTCAAGAAAG 240
Qy 2237 ATTTGATTTGCAATTTGACAAATCTTGTGTTCGGGAGAGGAGTTTCATCGTGTGGCAGCC 2296
Db 241 ATTTGATTTGCAATTTGACAAATCTTGTGTTCGGGAGAGGAGTTTCATCGTGTGGCAGCC 300
Qy 2297 TCAGCAAGCTCTCGGGGAGAGGGGCTCCAGCAGCGCATGTTCTTCTGTTCAAGACGTCC 2356
Db 301 TCAGCAAGCTCTCGGGGAGAGGGGCTCCAGCAGCGCATGTTCTTCTGTTCAAGACGTCC 360
Qy 2357 TGCTATACAGAGCGGGGGGCTGACGGGCTCCAACTAGTTTAAAGTCCACGGGAGCTCC 2416
Db 361 TGCTATACAGAGCGGGGGGCTGACGGGCTCCAACTAGTTTAAAGTCCACGGGAGCTCC 420
Qy 2417 CGCTCTATGCGATGACGATTTGAGGAGCGAAGCGAGTGGGGGTGCCCACTGCTGA 2476
Db 421 CGCTCTATGCGATGACGATTTGAGGAGCGAAGCGAGTGGGGGTGCCCACTGCTGA 480
Qy 2477 CCTCCGGGGGCGACGGGCACTCATCTGCTGGCGGCAAGTCTTCGCTCGAAGTGAAGA 2536
Db 481 CCTCCGGGGGCGACGGGCACTCATCTGCTGGCGGCAAGTCTTCGCTCGAAGTGAAGA 540
Qy 2537 AGTGGTTGAGGACATCCAGATGGCCATTGACCTGGGCGAGAGAGAGAGAGCGCCGCGCC 2596
Db 541 AGTGGTTGAGGACATCCAGATGGCCATTGACCTGGGCGAGAGAGAGAGAGCGCCGCGCC 600
Qy 2597 CTGAGTTCTCTGCGCAGCAGGCGCCCTTGACAACAAGTCCCTGATGAAGCCACCGGCTG 2656
Db 601 CTGAGTTCTCTGCGCAGCAGGCGCCCTTGACAACAAGTCCCTGATGAAGCCACCGGCTG 660
Qy 2657 ACCAGAGTCAAGAGATGACTTGAAGCGCTGCGGCACATGCTGGAAGCGCCAGCGCCGC 2716
Db 661 ACCAGAGTCAAGAGATGACTTGAAGCGCTGCGGCACATGCTGGAAGCGCCAGCGCCGC 720
Qy 2717 ACCGCGGCAACAAATGTTGACGTGTGCTGGACCGCAACACAGCGTCTCCATGCTGG 2776
Db 721 ACCGCGGCAACAAATGTTGACGTGTGCTGGACCGCAACACAGCGTCTCCATGCTGG 780
Qy 2777 ACTTCAGCATCGAGTGAAGATCACTGTCTGAAAACCTGCTGAGGAAATTCAAAAACA 2836
Db 781 ACTTCAGCATCGAGTGAAGATCACTGTCTGAAAACCTGCTGAGGAAATTCAAAAACA 840
Qy 2837 GCAAC-GGGTGGCAGAAAGCTGTGGTGTGTTCACAAATCTTCGCGCTG-----TTCTTCT 2890
Db 841 GCAACGGGGTGGCAGAAAGCTGTGGTGTGTTCACAAATCTTCGCGCTGTTTCTTCT 900
Qy 2891 ACAATACACACGAGCAAT---CATCCCTTGCCAGCTGCTCTG 2934

Db 901 ACBAATCACACGAGGACATTCAATTCCTGCGCAGCCTGCTG 947

RESULT 12
BX354920/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BX354920 1072 bp mRNA linear EST 23-Apr-2004
BX354920 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC027YA10 3-PRIME, mRNA sequence.
BX354920 GI:46552853
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1072)
Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30381927.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7229.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DC027BA0SNP1&c=7229.f.
Location/Qualifiers
1..1072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC027YA10"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source

Query Match 27.2%; Score 852.4; DB 4; Length 1072;
Best Local Similarity 98.6%; Pred. No. 5.5e-214;
Matches 897; Conservative 4; Mismatches 5; Indels 4; Gaps 4;

ORIGIN

2226 ACTCAAGAAAGATTTGATGGCATTCATCTTTGCTTCGGGAAGGAGTTTCATCCG 2285
1002 ACTCAAAAAGATTATTT-GCATTTTCAATCTTGTGTTT-CGGGAAAGAGTTTCATCCG 945

2286 TCTGGGAGGCTCAGCAAGCTCTCGGGGAAGGGGCTCAGACAGGATGTTCTTCGTT 2245
944 TCTGGGAGGCTCAGCAAGCTCTCGGGGAAGGGGCTCAGACAGGATGTTCTTCGTT 886

2346 CAACGAGCTCTCTATACAGAGCCGGGGGCTGACGGCTCCATCAGTTTAAAGTCCA 2405
885 CAACGAGCTCTCTATACAGAGCCGGGGGCTGACGGCTCCATCAGTTTAAAGTCCA 826

2406 CGGGCAGCTCCCGCTCTATAGCATGATGATTAGAGAGGAAAGAGATGGGGGTTGCC 2465
825 CGGGCAGCTCCCGCTCTATAGCATGATGATTAGAGAGGAAAGAGATGGGGGTTGCC 767

2466 CCACTGCTGACCTCCGGGGGCAAGGGGAGTCCATCATGTTGGCGCCAGTTCTGGGTC 2525
766 CCACTGCTGACCTCCGGGGGCAAGGGGAGTCCATCATGTTGGCGCCAGTTCTGGGTC 707

2526 CGAGATGAGAGAGTGGTTGAGGACATCCAGATGGCCATTGACCTGGCGGAGAGAGCAG 2585

Db 706 CGAGATGAGAGAGTGGTTGAGGACATCCAGATGGCATTTGACCTGGCGGAGAGAGCAG 647

QY 2586 CAGCCCCGCTCTGAGTTCTTGCCAGCAGCCCCCTTGACAAACAAGTCCCCCTGATTAAGC 2645
Db 646 CAGCCCCGCTCTGAGTTCTTGCCAGCAGCCCCCTTGACAAACAAGTCCCCCTGATTAAGC 587

QY 2646 CAGCGGCGTGAACAGAGATCAGAGATGACCTGAGCGGCTCGCGCAGATCGCTGAGAGG 2705
Db 586 CAGCGGCGTGAACAGAGATCAGAGATGACCTGAGCGGCTCGCGCAGATCGCTGAGAGG 527

QY 2706 CAGGCGCGGCAACCGGCGCAACAATGGTGACCTGTGCTGGACCGCAACACGCGT 2765
Db 526 CAGGCGCGGCAACCGGCGCAACAATGGTGACCTGTGCTGGACCGCAACACGCGT 467

QY 2766 CTCATGTTGACCTTCAGATGCGAGTGGAGATCATGTTGCTGGAACCTGCTGAGGA 2825
Db 466 CTCATGTTGACCTTCAGATGCGAGTGGAGATCATGTTGCTGGAACCTGCTGAGGA 407

QY 2826 ATTCAAAACAGCAACGGGTGGCAGAGAGCTGTGGGTGTTCAACAACCTTGCTGCTT 2885
Db 406 ATTCAAAACAGCAACGGGTGGCAGAGAGCTGTGGGTGTTCAACAACCTTGCTGCTT 347

QY 2886 CTTCTACAAATGACACACAGACCAATCATCCCTTGCAGCCTGCTTGCTGCTGCTT 2945
Db 346 CTTCTACAAATGACACACAGACCAATCATCCCTTGCAGCCTGCTTGCTGCTGCTT 287

QY 2946 GCTCACCAATCCCTCTGATTCGGAACATCCAGAAAGCTACGTTTGAAGTGAAGT 3005
Db 286 GCTCACCAATCCCTCTGATTCGGAACATCCAGAAAGCTACGTTTGAAGTGAAGT 227

QY 3006 CAAATCCAGCTCTACTACTTTCAGGGCGGAAAGCGAGTACAGTTTGAAGTGAAGTGA 3065
Db 226 CAAATCCAGCTCTACTACTTTCAGGGCGGAAAGCGAGTACAGTTTGAAGTGAAGTGA 167

QY 3066 AGTATCCGCAAGTCCACAGCTTGTGCTCGGACCCACGTTTGAAGTGAAGTGA 3125
Db 166 AGTATCCGCAAGTCCACAGCTTGTGCTCGGACCCACGTTTGAAGTGAAGTGAAGTGA 107

QY 3126 TCTTGTGTTAT 3135
Db 106 TCTTGTGTTAT 97

RESULT 13
AY415515
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AY415515 3165 bp DNA linear GSS 17-DEC-2003
AY415515
AY415515
AY415515.1 GI:39771474
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 3165)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS

Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 3165)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES
Source
1. .3165
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>3165
/gene="PARP2"
/locus_tag="HCM5583"

ORIGIN

Query Match 26.9%; Score 842.4; DB 14; Length 3165;
Best Local Similarity 51.9%; Pred. No. 3.3e-211;
Matches 1606; Conservative 0; Mismatches 1444; Indels 45; Gaps 7;

QY 21 GCCGACCCAGATCAAGACTGGGGGCCCGGAAATTCGGGATCAGTACCTTGAACG 80
DB 33 GCAGACTGCAAGGATGGCTTGGGTGCCAGAACCCCTGGGAGTTAGACCTTGAGCC 92
QY 81 TGGACAGAGCCGCCCAACACCTTCAGGAAACTGGTCCATTCAAATCCAGATGCT 140
DB 93 TGGGCAAGACTCTTGGCCAGAAATGCAAGAGACACCTGCACCTCAGAGTAAAGCTGCT 152
QY 141 GGAATGACCCAGGAGGCAATTGGAAGTTCCAAAGAGCTCCGGGAAAGGTCGCTGGA 200
DB 153 GGAACACACATGGAAATATTTGACATTGAGCCTTAATTCGATGCGCAGGATTAAGTAC 212
QY 201 TGCAGTTGGCAACCACTCAACCTCGTGAAGTGACTATTTTGGCTCGAGTTCTCTGA 260
DB 213 ACAAGTGGAAAGGTTTAAACCTGTGAATGTGACTACTGGGATGAGTTTCAAAA 272
QY 261 TCACAAAAGATCAAGCTGTGGCTGATCTCTTAAACCAATTGTGAAAGATTAGAG 320
DB 273 TACTCACTCTACTGGAATTTGGCTTGAACCTATGAACCACTATTAGCAAAATACGAG 332
QY 321 GCCAAGCAAGTGTGTGTAAGTTGTGTAAGTTCTTCCGCTGACACACAGACT 380
DB 333 GCCAAGAAATGTGTGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 392
QY 381 CCAAGAAAGATCAAGAGTACCTGTTCCGCTGCAGGTGAGAGCAAGCACTTGGCTCAGG 440
DB 393 ACAAGAAAGATATACAGATATCTTGTGCTTGCATTAAGAGACCTGCTGAGAA 452
QY 441 CAGGTGAGCTGTATGACACACAGCGAGCTCTTGAATTTACACATTGTGCAATCTGA 500
DB 453 GCGTTTGACCTGTGTCACACACAGCGGCTTCTCAGCTCCCATCTCTGCACTGCGGA 512
QY 501 GATTGGGATTTTGAAGACCTTGAACAGAGCACTTGAAGCAAAAATAATATCATACC 560
DB 513 AATATGAGATTAAGATGAACGCTGACCGAGACACCTCAAGTGAACAGATTTGCTC 572
QY 561 TCAGCAAGACGACTAGAGGACAAATCGTGAATTTCAACATTAACCAATTGGACAAAC 620
DB 573 TGGCCAGACACATCGCTTGAAGATATCTAGATTTCACTCAGAGACATGCGGCCAGAC 632
QY 621 ACCAGCAAGATCAAGATTTCAAGCTCTTGAAGATTCCTGCGCTAGAGATGTATGGAAT 680
DB 633 ACCTGCTGATCGGATTTCCAGGCTGCTGAAATTTGCTCGAAAGTTGAAATGTACGCGCAT 692
QY 681 CCGGTTGCAACCGGCGCAAGACAGGGAAGGACGAGATCAATCTGCGCTGCAACAC 740
DB 693 CAGATTTCAATATGCTTCTGACAGGGAAGAACCAAGATTTCAATGAGCAATTTCCCAT 752
QY 741 GGGAAATCTAGTGTTCAGGATTTCACTAAGATCAATGCTTCAACTGGGCAAGTGC 800
DB 753 GGGGTACTCGTGTTCAGNN 812
QY 801 GAAAGCTAGCTTCAAGAGGAAGCGCTTCTCATCAAGCTCGGCGCAGATGCAATAGTGC 860
DB 813 NNGACC 872

QY 861 GTACAGATACCTTGAATTCCTGATGCGCAGTGGGATTTCTGCAAGTCTTCTGAA 920
DB 873 TTACAGGACACATTTAGATTTTGTGGTAGTAGATGATGTAAAGTCTTCTGAA 932
QY 921 AATCTGTGTGAACATCATGCTCTTTTACACTTTTGAAGAGCCCAACCAAGCCCAA 980
DB 933 GATTGTGTAGATATCAACCTTTTATGACTTTTGAACCAACTTAAGCAAAAGCAAA 992
QY 981 GCCGCTCTTTTACCGGGGGTCATCATTTGGTTCAGTGTGCGACTGAGAGAGCT 1040
DB 993 AGCGCTTCTTCAAGCGGGGCTCTCTTCAATTAAGNNNNNNNNNNNNNNNNNNNNNN 1052
QY 1041 TCTGACTATGTTAAAGAAAGAGACATAAGAGTGCAGTTGAAAGAGACAGCAA 1100
DB 1053 NNN 1112
QY 1101 GATTCAATTCATCGGAGCCTTCTTCAAGCTTACAGAACTGAATTCGAAGTCTGGA 1160
DB 1113 GACCCACAGCTC-----GTTGAGCTCTGACTGACAGACTTACCAAA 1154
QY 1161 GCAGTCTCAGACAGACACAGCCTTACATTGAGAGAGTGGCCGAATCTCCAGGGGCCA 1220
DB 1155 ACAGAGCATCTATTCCTGAGGATTTGAGAGCTCTGCTCCCATCTTCAAGGAATGC 1214
QY 1221 GAGCTGCGCGAGAGAAAGGAAACGAAAGTTTCGCGCGGAGACCGGGAGTCGACCCGAG 1280
DB 1215 CTTTACTGCTCTCTCCCTCACTCTGCTCCCTCTGCGCTGCGAGAGTTTAAAGACAG 1274
QY 1281 CCTGCGCGAGAGAGAGCCCGGGGTAAACAAGACGAGCGAGCGCTCTGCGGCC 1340
DB 1275 CAGAGCTCCTCAGAGATCCAGAGTTTCTTACGTCAAGAGTCCAGCTCAGAGAGCG 1334
QY 1341 CAGGAGGAAGAGAGAGAG---TGTGAAGATAGAGCCACAGAGTAACTCAGCC 1397
DB 1335 CAGTGAAGCAGTGTGAGAGGCCCAACACATCGGCTCAGCCCTCGAGCCCGCC 1394
QY 1398 CCGCAGCCAGAGACAGAGCTCCCTGACTGGCAGTCTCCTTCCGAGCTGTCTGTGAA 1457
DB 1395 ACTTCAGCTGTGTCCAGGCTTTCAGAGAAAGTCTTCCAGCTTCTCCCTCCAGCGGAA 1454
QY 1458 CTGCAAGGGGAGAGTGGCCCTGCAACGTGACCTTGTCTCCCACTGAGCCCGAC 1517
DB 1455 GAGCCCTGAGTCTGAGCCCTGATTTCAAGTGCCT-----TGAGCCCACTGA 1505
QY 1518 CAAGCAGGCTCTCTCTTGAATCAGCCGCTGCTGAATGACACAGGCTCCCGCCGAGCGA 1577
DB 1506 ACAAGGCTCATCCCACTCTGAGCCTGTCTCAAGTGAAGTGTGGCGAGCCGGAATGA 1565
QY 1578 CGATGAGGATGAGGCGCGAGAGAGATTCCTCACTGATTAAGCGTACTTAGCTAA 1637
DB 1566 CTGCGAG--GAGCCAGACACAAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1622
QY 1638 GGAAGTGTCTACACAGAGGAAATATCTGAAGATCTCGAAGTTATCATCTTGTGTT 1697
DB 1623 NNN 1682
QY 1698 TCAGAGCAGTGAAG 1757
DB 1683 CCGCAGGCGAGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1742
QY 1758 GAATTTTGAACCTTTCACAAATTTTCACTAATTTTCTCAAGAAATTTGAGCAAGACT 1817
DB 1743 CAACATCATATCATCTATGAGTTTCACAGAGAGCTTCTGCGAGAGTGAAGAGAGCT 1802
QY 1818 TGCCTGTGGAAGCGCGCTCAATTCCTCAAG--ATTACCAAGAAATCGGCGA 1874
DB 1803 GGCACCTCTGGAGAGGCGCTCCCAAGCCCAACAAAGGCAATCAACAGAAATCGGGGA 1862
QY 1875 TGTCTATGTAAGAAATTCAGGAGCATTAAGCACTGCGGCTCACCTGTGAGAGACAG 1934
DB 1863 CATCTGCTCAGGAACATGCGCAATTAAGAGGTTTACAGCTTCTCAAGAGATGA 1922
QY 1935 CGAGGCTTGAAGGCGCTGAGAGATGAATCAAGAGCTCCGCGGCTGAGAACTTCTG 1994

```

Db      1923 CGAGGTCCTTAACGAACTGGAAAAAGGCTACCAACGCTGTAAAGATGGAGGACAGTGA 1982
Qy      1995 CAGAGACTTTGAGCTGAGAAAGGTGTGTTCTTACCGCTCAACACCTCTCTCGGAGCC 2054
Db      1983 CAAGGAGTTTGAAGCTGAGAAAGGTCTGTCTACTTGTCTTCAACAGCTTCTGTGTGAAGCC 2042
Qy      2055 ACTGCACCGGCTCATGCTACCAAGCAGGTCTGAGCGGCTGTGCAACCAACCGGCGC 2114
Db      2043 CATTCAGCGGCTGCTGTGACTACCGCCCTGTGCTGCGCGCTTATGCGAGCATTAACAGCCC 2102
Qy      2115 GAGCCACGCGCACTTCAAGGACTGCGAGCGCTTTGGCAGAGATCACGAGATGCTGCG 2174
Db      2103 CGGGACACATGACTAGCGCTGAGCCATGACGCCCTGAAGCATCAACAGGTGACACAC 2162
Qy      2175 ACAGCTCCACGGTACCATGATCAAGATGAGAAATTTCCAAACCTGACAGCAATCTCAAGAA 2234
Db      2163 CACACTACAGCACAATCTCATCCGGCTGAGAACTTGCAAAAGCTTAAACGAGCTGACGCG 2222
Qy      2235 AGATTGATTGGCAATTGACAAATCTTGTGTTCCGGAAAGGAGTTCAATCCGTCTGGGACG 2294
Db      2223 GGAAGCTGTGGGCAATGAGAACTCATGTCTCTGGGAGGAGTTCAATCCGTGAGGCGTG 2282
Qy      2295 CCTCAGCAAGCTCTCGGGAAAGGGGCTCCAGACAGCGCATTTCTCTGTTCACGACGT 2354
Db      2283 CTTTCAAGACTCACCAAGAAAGGGCTGACAGAGAGATGTTTTTCTGTTCAGATAT 2342
Qy      2355 CCTGCTATACAGACCGCGGGGCTGACGCGCTCAATCAATTTAAAGTCCACGGGACGT 2414
Db      2343 GTTGTGTATACAGCAAGAAAGAGTGTGAGGAGCAAGCACTTCGGATCGGGGCTTCT 2402
Qy      2415 CCGGCTCTATGATGATGATGATGAGAGAGCAAGCAAGTGGGGGTGCGCCACTGCT 2474
Db      2403 TCCCTCTCAAGGCAATGTGTGTGAGAAAGATGATTAAGATGTGTCTTCCACATGTTT 2462
Qy      2475 GACCTTCGGGGCCAGCGGCAATCATGTGTGCGCGCAAGTTCTGTGCTCCAGATGA 2534
Db      2463 CACCATCTACGCGGCTCAGAAACCAATCGTGTGAGCAGCAGCACTCGGCTGAGAAAG 2522
Qy      2535 GAATGTGGTGAAGACATCCAGATGGCCATTTGACCTGGCGGAGAAAGACGACGCCCCG 2594
Db      2523 GAAGTGAATGCTGAACTGAACTCCGATTCAGACGACCAAGATGGCGGTGACACGGC 2582
Qy      2595 CCCTGAGTTCTGTGCGCAGCA-----GCCCCCTGCAACAAAGTCCCTGATGAACCCAC 2648
Db      2583 CCTGTGACGTGCGCGGCACTGTGTGCACTGTCCCCCAAGATCCCCCAAGAGGATTC 2642
Qy      2649 CGCGGCTGACAGAGATCAGAGATGACTGAGCGGCTCGCGCAATCGCTGAGCGCCA 2708
Db      2643 TCTG---GAGCAGAGATCAGAAAGATGCTCGGGGTGTCCGACAGCTCCCTGAGAGGGCA 2699
Qy      2709 GGGCCCGGCAACCGGGGCAACAAATGTGTGCGGTGTGCAACCGCAACGACGCGTTC 2768
Db      2700 TGCGCAGCAGCGGGCAACCAACCAATGCACTGTGTGTGATCCGGAACCAAGCGTGTTC 2759
Qy      2769 CATGTGTGACTTGTAGCATCGCACTGAGAAATCAAGTTGTCTGAGAAACCTGTGAGAAAT 2828
Db      2760 CAGGGCAGACCAAGTGTGAGCTGTGCGAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2819
Qy      2829 CAAAAACAGCAACGGGTGCAAGAGCTGTGTGTGTTCACAACTTCTGCTGTCTT 2888
Db      2820 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2879
Qy      2889 CTACAAATCACAACAGCAATATATCCCTTTCAGAGCTGCTCTGTGCGCTACTCGCT 2948
Db      2880 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2939
Qy      2949 CACCATCCCTCTGTAGTCCGAGAAATCCAGAAAGCTACGTGTTCAGAGCTTCA 3008
Db      2940 GAGCATCTCCAGGAGGCGATGAGCAATACAAAGACTATGTTTCAAGTCCAGTTCAA 2999
Qy      3009 GTCCCACTTACTTCAAGGGCGAAGACGATACCTTGAAAGGTGATGAGAT 3068

```

```

Db      3000 ATCCACGCTCTACTTCTTCCGGGCTGAGAGCAAGTACATTTGAAAGGTGATGAGGT 3059
Qy      3069 GATCCGAGTGGCACACAGCTGTGCTCGGACGCC 3103
Db      3060 GATTCAGGGGGCCAGAGCTCAGCGGAGGCGCC 3094

RESULT 14
BUS39080
LOCUS      BUS39080          959 bp      mRNA      linear      EST 13-SEP-2002
DEFINITION AGENCOURT_10215257 NIH_MGC_107 Homo sapiens cDNA clone
IMAGE:6569730 5', mRNA sequence.
ACCESSION  BUS39080
VERSION     BUS39080.1  GI:22849521
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homidae; Homo.
REFERENCE   1 (bases 1 to 959)
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            JOURNAL   Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            COMMENT   Email: cgaabs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            plate: LLCM2757 row: h column: 18
            High quality sequence stop: 706.
FEATURES
            source
            1..959
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6569730"
            /tissue_type="adenocarcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_1ib="NIH MGC 107"
            /note="Organ: breast; Vector: pOTB7; Site 1: EcoRI;
            Site 2: XhoI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCACGAG(G). Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC Library."
ORIGIN
Query Match      26.8%; Score 839.8; DB 3; Length 959;
Best Local Similarity 97.8%; Pred. No. 1.2e-210;
Matches 861; Conservative 0; Mismatches 18; Indels 1; Gaps 1;
Qy      811 TTCAAGAGAACGCTTTTCTCATCAAGCTCGGCGCAAGATGCCAATGTGCGTACCGAGAT 870
Db      1 TTCCGACAGAGCGCTTCTCTCAAGCTCCGCGCAGATGCCAATGTGCGTACCGAGAT 60
Qy      871 ACCTTGAATTCCTGATGCGAGTGGGATTTCTGCAAGTCTTCTGAAAAATCTGTGTT 930
Db      61 ACCTTGAATTCCTGATGCGAGTGGGATTTCTGCAAGTCTTCTGAAAAATCTGTGTT 120
Qy      931 GAAATCATGCTCTTTTGAAGAGCCCAAGCCCAAGCCCAAGCCGCTCTC 990
Db      121 GAAATCATGCTCTTTTGAAGAGCCCAAGCCCAAGCCCAAGCCGCTCTC 180
Qy      991 TTTAGCCGGGGGTATCATTTTGGTTCAAGTGTGCGACTCAAGAGAGGTTCTCACTAT 1050
Db      181 TTTAGCCGGGGGTATCATTTTGGTTCAAGTGTGCGACTCAAGAGAGGTTCTCACTAT 240

```

QY 1051 GTTAAAGAGAGACATTAAGATGCAATTTGAAAGAGACACAGAAATTCATTCT 1110
DB 241 GTTAAAGAGAGACATTAAGATGCAATTTGAAAGAGACACAGAAATTCATTCT 300
QY 1111 ATCCGAGCCTTGTCTTACAGCCTTACAGAACTGAATTCGAAGTGTGAGCACTCTCAG 1170
DB 301 ATCCGAGCCTTGTCTTACAGCCTTACAGAACTGAATTCGAAGTGTGAGCACTCTCAG 360
QY 1171 CAGAGCACACGCTTACATTTTGAGAGAGTCCCAATCTCCAGGAGGCTGAGCTCCCG 1230
DB 361 CAGAGCACACGCTTACATTTTGAGAGAGTCCCAATCTCCAGGAGGCTGAGCTCCCG 420
QY 1231 CGAGAGAAAGAAACGAAAGTTCCTCGCGGAGAGCCGGGTGACACCCGAGCCCTGCGCCG 1230
DB 421 CGAGAGAAAGAAACGAAAGTTCCTCGCGGAGAGCCGGGTGACACCCGAGCCCTGCGCCG 480
QY 1291 AGGAGAAAGCCCGCGGTAAACAGACAGCGAGCCGCTCGCGCCCAAGAGAA 1350
DB 481 AGGAGAAAGCCCGCGGTAAACAGACAGCGAGCCGCTCGCGCCCAAGAGAA 540
QY 1351 GAGAGAGAGTCTTAAAGATAGAGCCAGACAGTAACCTCAGCCCGCAGCCAGC 1410
DB 541 GAGAGAGAGTCTTAAAGATAGAGCCAGACAGTAACCTCAGCCCGCAGCCAGC 600
QY 1411 AAGAGCTCCCTGACCTGAGCTTCACTTCCGAGCTGTGTGAATCTGCAAGGGGGA 1470
DB 601 AAGAGCTCCCTGACCTGAGCTTCACTTCCGAGCTGTGTGAATCTGCAAGGGGGA 660
QY 1471 GTGAGCCCTGCAACAGTACCTTGTCTCCCACTGAGCCCGCAGACCAAGAGGCTCT 1530
DB 661 GTGAGCCCTGCAACAGTACCTTGTCTCCCACTGAGCCCGCAGACCAAGAGGCTCT 720
QY 1531 CCTTGTACAGCCGCTGCTGAATGACCAAGCCCTGCGCCGAGAGATGAGATGAG 1590
DB 721 CCTTGTACAGCCGCTGCTGAATGACCAAGCCCTGCGCCGAGAGATGAGATGAG 780
QY 1591 GGCGGAGAGAAAGATTCCTCACTGAATGAAGCTTCAATGATTAAGAAAGTCTTACC 1650
DB 781 GGCGGAGAGAAAGATTCCTCACTGAATGAAGCTTCAATGATTAAGAAAGTCTTACC 839
QY 1651 ACCGAGCAATATCTGAAGATCTCGAAGTTTCACTT 1690
DB 840 ACCGAGCAATATCTGAAGATCTCGAAGTTTCACTT 879

RESULT 15
BX350128/c 857 bp mRNA linear EST 08-APR-2004
LOCUS BX350128 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1043YD12 3-PRIME, mRNA Sequence.
ACCESSION BX350128
VERSION BX350128.1 GI:30367428
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 857)
L1.W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7229.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0D10222D02_CS02028_1&c=7229.f

FEATURES
source Location/Qualifiers

1..857
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1043YD12"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 26.6%; Score 834; DB 4; Length 857;
Best Local Similarity 99.8%; Pred. No. 3.9e-209;
Matches 856; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1689 TTGCTGTTTCAAGACACAGTGAAGCAAAAGAGCGCCATCCGAGACCTGAAAGTCT 1748
DB 856 TTGCTGTTTCAAGACACAGTGAAGCAAAAGAGCGCCATCCGAGACCTGAAAGTCT 797
QY 1749 CATATTCGCAATTTTGAACCTTGTGCAAAATTTCAATATTTTCTCAAGAAATTTGA 1808
DB 796 CATATTCGCAATTTTGAACCTTGTGCAAAATTTCAATATTTTCTCAAGAAATTTGA 737
QY 1809 GCAACGACTTGCCTGTGAGAGAGCGCTCAAAATGCCAATCAGAGATTACCAAGAAAT 1868
DB 736 GCAACGACTTGCCTGTGAGAGAGCGCTCAAAATGCCAATCAGAGATTACCAAGAAAT 678
QY 1869 CGCGATGTATCTGTAAGAAATTCAGGCGATGAGACACCTGCGGCTCACTGTGAA 1928
DB 677 CGCGATGTATCTGTAAGAAATTCAGGCGATGAGACACCTGCGGCTCACTGTGAA 618
QY 1929 GCAACGAGAGCCCTTGAAGAGCCCTTGAAGATGAAATCAAGTCCCGGAGGCTGAGAA 1988
DB 617 GCAACGAGAGCCCTTGAAGAGCCCTTGAAGATGAAATCAAGTCCCGGAGGCTGAGAA 558
QY 1989 CTTCTGAGAGACTTTGAGCTGCAAGAGTGTGTACTTACCGCTCAACACTTCTCTCT 2048
DB 557 CTTCTGAGAGACTTTGAGCTGCAAGAGTGTGTACTTACCGCTCAACACTTCTCTCT 498
QY 2049 GCGGCCACTGCACCGGCTCATGCACTAACAGAGTCTTGAAGCGGCTGTGCAACCA 2108
DB 497 GCGGCCACTGCACCGGCTCATGCACTAACAGAGTCTTGAAGCGGCTGTGCAACCA 438
QY 2109 CCGCGGAGGCAAGCGGCACTTCAAGGAGCTGCGGAGCGCTTGGCAGAGATCAGGAGAT 2168
DB 437 CCGCGGAGGCAAGCGGCACTTCAAGGAGCTGCGGAGCGCTTGGCAGAGATCAGGAGAT 378
QY 2169 GGTGACACAGCTCCACGCTACAGTGAATCAAGATGAGAAATTTCCAGAGCTGACAGAACT 2228
DB 377 GGTGACACAGCTCCACGCTACAGTGAATCAAGATGAGAAATTTCCAGAGCTGACAGAACT 318
QY 2229 CAAGAAAGATTTGATTTGCAATTCATTTGTTTCCGGAAGGAGTTCAATCCGCTCT 2288
DB 317 CAAGAAAGATTTGATTTGCAATTCATTTGTTTCCGGAAGGAGTTCAATCCGCTCT 258
QY 2289 GGGCAGCTCAGAGAGCTTCGGGGAAGGGGCTCCAGACAGGATGTTCTTCCGTTCAA 2348
DB 257 GGGCAGCTCAGAGAGCTTCGGGGAAGGGGCTCCAGACAGGATGTTCTTCCGTTCAA 198
QY 2349 CGAGCTCTGCTTATACAGAGCCGGGAGCTGACGCGCTCCAAATCAGTTTAAATCCACGG 2408
DB 197 CGAGCTCTGCTTATACAGAGCCGGGAGCTGACGCGCTCCAAATCAGTTTAAATCCACGG 138
QY 2409 GCAAGCTCCGCTCTATGACATGACAGATTGAGAGAGCAAGAGTGGGGGTGCCCA 2468
DB 137 GCAAGCTCCGCTCTATGACATGACAGATTGAGAGAGCAAGAGTGGGGGTGCCCA 78

QY 2469 CTGCTGACCCCTCCGGGGCCAGCGGCAATCATCGTGGCCGCAATTCTCGATCCGA 2528
DB 77 CTGCTGACCCCTCCGGGGCCAGCGCAATCATCATCTGTTGCCCCCAATTCTCGATCCGA 18
QY 2529 GATGAGAAAGTGGGTTGA 2546
DB 17 GA-GAAGAGTGGGTTGA 1

Search completed: May 22, 2006, 08:59:33
Job time : 13995.1 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2006, 00:49:20 ; Search time 542.44 Seconds
(without alignments)
10813.957 Million cell updates/sec

Title: US-09-555-342b-1_COPY_49_3183

Perfect score: 3135
Sequence: 1 atggggaagatagagcagag.....acaaagagctctctgtatc 3135

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /EMC_Celerra_SIDS3/prodata/2/ina/1_COMB.seq:*
- 2: /EMC_Celerra_SIDS3/prodata/2/ina/5_COMB.seq:*
- 3: /EMC_Celerra_SIDS3/prodata/2/ina/6A_COMB.seq:*
- 4: /EMC_Celerra_SIDS3/prodata/2/ina/6B_COMB.seq:*
- 5: /EMC_Celerra_SIDS3/prodata/2/ina/7_COMB.seq:*
- 6: /EMC_Celerra_SIDS3/prodata/2/ina/H_COMB.seq:*
- 7: /EMC_Celerra_SIDS3/prodata/2/ina/PC/US_COMB.seq:*
- 8: /EMC_Celerra_SIDS3/prodata/2/ina/PP_COMB.seq:*
- 9: /EMC_Celerra_SIDS3/prodata/2/ina/RE_COMB.seq:*
- 10: /EMC_Celerra_SIDS3/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	204.2	6.5	310	US-09-621-976-8193	Sequence 8193, Ap
2	192.4	6.1	2853	US-09-949-016-1933	Sequence 1933, Ap
3	180.4	5.8	6263	US-09-664-958-5	Sequence 5, Appl1
4	180.4	5.8	6268	US-09-566-921-57	Sequence 57, Appl1
5	169.6	5.4	4292	US-09-949-016-1488	Sequence 1488, Ap
6	169.6	5.4	4336	US-09-949-016-208	Sequence 208, Ap
7	165.8	5.3	576	US-09-270-767-1995	Sequence 1995, Ap
8	165.8	5.3	576	US-09-270-767-1727	Sequence 1727, A
9	144	4.6	3984	US-09-848-294-1	Sequence 1, Appl1
10	143.2	4.6	2872	US-09-906-779-3	Sequence 3, Appl1
11	142.4	4.5	2156	US-09-949-016-3593	Sequence 3593, Ap
12	133.6	4.3	3398	US-09-799-451-201	Sequence 201, Ap
13	106.4	3.4	449	US-09-270-767-1542	Sequence 1542, Ap
14	106.4	3.4	449	US-09-270-767-16824	Sequence 16824, A
15	105.2	3.4	184	US-09-621-976-15728	Sequence 15728, A
16	97.8	3.1	17098	US-09-949-016-15864	Sequence 15864, A
17	95	3.0	2493	US-10-104-047-41	Sequence 41, Appl1
18	90.6	2.9	573	US-09-270-767-165	Sequence 165, Appl
19	90.6	2.9	573	US-09-270-767-15447	Sequence 15447, A
20	65.2	2.1	3438	US-10-164-595-29	Sequence 29, Appl1
21	64	2.0	4272	US-09-949-016-5747	Sequence 5747, Ap
22	63.8	2.0	1626	US-09-620-312D-1033	Sequence 1033, Ap
23	63.8	2.0	1699	US-09-913-171A-26	Sequence 26, Appl

24	62.8	2.0	2730	US-09-774-528-175	Sequence 175, App
25	62.8	2.0	2730	US-10-120-988-175	Sequence 175, App
26	62.8	2.0	3503	US-10-094-749-560	Sequence 560, App
27	59.6	1.9	722	US-09-270-767-1861	Sequence 1861, App
28	59.6	1.9	722	US-09-270-767-17143	Sequence 17143, A
29	59.4	1.9	222	US-09-016-434-45	Sequence 45, Appl1
30	56.8	1.8	4080	US-08-446-345-35	Sequence 35, Appl1
31	56.8	1.8	4081	US-09-949-016-1292	Sequence 1292, Ap
32	53.8	1.7	7218	US-08-232-463-14	Sequence 14, Appl1
33	50.2	1.6	3044	US-09-814-915A-84	Sequence 84, Appl1
34	50.2	1.6	3044	US-09-880-107-3718	Sequence 3718, Ap
35	48.8	1.6	1308	US-09-252-991A-11936	Sequence 11936, A
36	48.8	1.6	1335	US-09-252-991A-12073	Sequence 12073, A
37	48.2	1.5	1816	US-09-620-312D-1439	Sequence 439, App
38	48.2	1.5	3415	US-09-902-540-5107	Sequence 5107, Ap
39	48.2	1.5	28058	US-09-902-540-1252	Sequence 1252, Ap
40	47.4	1.5	2400	US-08-930-001-1	Sequence 1, Appl1
41	47.4	1.5	2400	US-09-091-885-1	Sequence 8879, Ap
42	47	1.5	801	US-09-902-540-8879	Sequence 1071, Ap
43	47	1.5	12028	US-09-902-540-1071	Sequence 17489, A
44	47	1.5	54711	US-09-949-016-17489	Sequence 17489, A
45	46	1.5	235452	US-09-949-016-13675	Sequence 13675, A

ALIGNMENTS

```
RESULT 1
US-09-621-976-8193
; Sequence 8193, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8193
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8193

Query Match          6.5%; Score 204.2; DB 3; Length 310;
Best Local Similarity 98.1%; Pred. No. 1.8e+4;
Matches 203; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGGGAAATAGAGCAGAGGCCGACCCAGATCAAGCTGGGGCCCCGGAAATTGC 60
      |||
DB      104 ATGGGAAATAGAGCAGAGGCCGACCCAGATCAAGCTGGGGCCCCGGAAATTGC 163

QY      61  GGGATCAGTACCTTGGAGAGTGGAGCCGCCCAACCTTCAGAAATCGTG 120
      |||
DB      164 GGGATCAGTACCTTGGAGAGTGGAGCCGCCCAACCTTCAGAAATCGTG 223

QY      121 TCCATCAATCAGATGCTGATGACACCCAGAGCAATTTGAACTTCACAAAGACT 180
      |||
DB      224 TCCATCAATCAGATGCTGATGACACCCAGAGCAATTTGAACTTCACAAAGACT 283

QY      181 CTTGGAGAGTGTCTGTGATGCAATT 207
      |||
DB      284 CTTGGAGAGTGTCTGTGATGCAATT 310

RESULT 2
US-09-949-016-1933
; Sequence 1933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1933
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-1933

```

```

Query Match      6.1%; Score 192.4; DB 3; Length 2853;
Best Local Similarity 52.8%; Pred. No. 7.8e-41;
Matches 490; Conservative 0; Mismatches 426; Indels 12; Gaps 3;

```

```

Qy 117 CGTGTCCATCAAAATCCAGATGCTGGATGACACCCAGAGGCACTTTGAAATTCACAAAG 176
Db 784 CATGCATCGCAAGGTTCTTTGTTGATGACACAGTTTATGATGTTGTGTGAGAAACA 843
Qy 177 AGCTCTGGAGAGTCTGCTGATGATGATGCAACCTCAACCTGCTGAGAGTGA 236
Db 844 TGTCTAAGGACAAAGATTTGCTTAAAGAGATGTGAGACTTCACATCTTTGGAGAAAGA 903
Qy 237 CATTTTGGCCCTCGAATTTCTGATCAAAAAGATCAAGGTGTGCTGGATCTCTAAA 296
Db 904 CATTTTGTCTAGCAATTTGGATTAAGCACTTAAAGACATGCTGAGATCCGCCAA 963
Qy 297 ACCCATTTGTGAAACAGATTAGAAGGCCAAAGCAGTTGTGTTAAGTTTGTGTGAATT 356
Db 964 AGAAATTAAGAAAGCAAGTTCGTGG---TGTCCCTTGGAAATTTTACATTTAATGAATTT 1020
Qy 357 CTTTCGGCTGACACACACAACTCCAAAGAACTCAAGATACCTGTTGGCGTGA 416
Db 1021 TTATCCACCTGACCCAGACAGTTAACAGAGACATACAGATATTTATTTATGCTTCA 1080
Qy 417 GGTGAAGCAGGACTTGGCTCAAGGCAAGTTGAGTGTATGACACAGGCGAGCTCTT 476
Db 1081 GCTTCGGCAGGACATAGTTGACGAGAGCTGCTGCTGCTGCTTCTTGAACCTTGAATTAAT 1140
Qy 477 GATTTCACATTTGTCATCTGAGATTGGGATTTTGTATGAAGCCCTTGACAGAGACA 536
Db 1141 AGGTTCTTACACATCTCAAGTCTGAAGTGGGAGCTAGACCCGAACTCATGCGCTGA 1200
Qy 537 CTTAGCAAAAATTAATACATACCC-----TCAGCAAGCAGCTAGAGACAAATTCGT 590
Db 1201 TTATGTATGATGATTTTAACTGGCCCCGATCAGACCAAGAACTTGAAGAGAGGTCAAT 1260
Qy 591 GGAATTTGCAATTAACCAATTTGACAAACACAGAGATCAAGATTTCCAGTCTCTAGA 650
Db 1261 GGAATGCTAATGATCAATCAAGGCTCAATGATCCAGCTCAGGCTGAGTCTGAGTTCTTGA 1320
Qy 651 GATTGCCGCTGGCTAGAGATGATGAATTCGCGTTGACCCGCGCAAGAGACAGGAAAG 710
Db 1321 GATGTCACAAAAGTTGTCTATGATGAGATGATCTTCAATTAAGCAAGAGACTTGAAG 1380
Qy 711 CACGAAGATCAATCTGCGCGTTGCAACACGGAATTTCTAGTCTTCAAGGTTTCACTAA 770
Db 1381 AGTAGATATCATCTCTAGGTCTGCTCTGATGCGCTTCTGTTTACAAAAGATTAAGCTGAG 1440
Qy 771 GATCAATGCTTCAATCGGGGCAAGGTCGGAAGTGAAGTCAAGAGAGAGGCTTTCT 830
Db 1441 AATTAACCGCTTCCCTTGGCCCAAGTGTGAAGATTTCTTATTAACGTAAGTACTTTT 1500
Qy 831 CATCAAGCTCCGGCAGATGCAATGATGCTACAGATACCTTGAATTCCTGATGCG 890

```

```

Db 1501 CATCAAGATTCGGCCCTGGAGAGCAAGACGATATGAAGTACATCGATTCAACTCC 1560
Qy 891 CAGTGGGATTTTCTGCAAGCTCTTCTGAAAAATCTGTGTGAACATCATGCTTCTTAG 950
Db 1561 CAGTTACCGAGCAGCTAAGAAATTTATGAAAGTCTGTGTGAACATCACACGTTTTTCAG 1620
Qy 951 ACTTTTGAAGAGCCCAACCAAGCCGAGCCGCTCTTTAGCCGGGGTCAATCT 1010
Db 1621 A---TTGACATCTTACAGACACCATTCCTCAAAAGCAAAATTTCTTGGCTAGATCCAAAT 1677
Qy 1011 TCGTTCACTGCTGCACTCAGAACAG 1038
Db 1678 TCGATACAGTGGCCGCACTCAAGCTCAG 1705

```

```

RESULT 3
US-09-664-958-5
; Sequence 5, Application US/09664958
; Patent No. 6916912
; GENERAL INFORMATION:
; APPLICANT: Trakht, Ilya
; APPLICANT: Canfield, Robert
; APPLICANT: Kalantarov, Gary
; APPLICANT: Rudchenko, Sergei
; TITLE OF INVENTION: No. 6916912e1 Tumor-Associated Marker
; FILE REFERENCE: 0575/60240
; CURRENT APPLICATION NUMBER: US/09/664,958
; PRIOR FILING DATE: 2000-09-18
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patencin version 3.1
; SEQ ID NO 5
; LENGTH: 6263
; TYPE: DNA
; ORGANISM: Human
; US-09-664-958-5

```

```

Query Match      5.8%; Score 180.4; DB 3; Length 6263;
Best Local Similarity 52.9%; Pred. No. 1.9e-37;
Matches 485; Conservative 0; Mismatches 416; Indels 15; Gaps 4;

```

```

Qy 137 TCGTGAATGACACCAAGAGCAATTTGAATTCACAAAGACTCTGGAAAGTGTGTC 196
Db 469 TCGTGAATGCTCGGAGTATGATGTAGTGAAGAAACATGGCCGGGCGAGGTGCTGT 528
Qy 197 TGGATGCAATTTGCAACCACTCAACCTGCTGGAAGGTGATTTGGCCCTGAGTTTC 256
Db 529 TTGACCTGCTGTGAACCACTCAACCTCTAGAGAAAGACTAATCTGCGCTGACCTTCT 588
Qy 257 CTGATCACAAAAAGATCAACGCTGCTGATCTCTTAAACCAATTTGAAACAGATTA 316
Db 589 GTGATGTCAGACGCAAGAAAGTCTGCTGACCTCTCCAGAGATCAAGAAAGAGATCC 648
Qy 317 GAAGGCCAAAGCAGTGTGTTAAGTTGTGTGAATTTCTTCCGCTGACACACAC 376
Db 649 GGAG---TAGCCCTGGAATTTTGCTTACAGTCAAGTCTAACCCGCTGATCTCCGCC 705
Qy 377 AACTCCAGAAAGAACTCAAGATGATCTGTGCGCTGAGGTGAAGAGGACTTGGCTC 436
Db 706 AGCTGACGAAAGACATCAAGATGATCTACCTGTGCTGAGCTGCGGCGAGCATATCA 765
Qy 437 AAGCAGGTTGACGTTAATGACACCAAGCAGCTCTCTTGAATTTCAACATTTGTCAT 496
Db 766 GGGGCGGCTGCAAGTCTTGTGACGCAATGCTCTGAGGCTCTCAAGCTGTGACAG 825
Qy 497 CTGAGATTTGGGATTTTGTATGAAGCTTGGACAGAGAGCACTTGAACAAAATTAATCA 556
Db 826 CTGAGCTGGTGAATGATGCTGAGAGCATGTGGGCAATGATGTCAGCGAGCTCCGCT 885
Qy 557 TAC-----TCAGCAAGCAGCACTAGAGGACAAATGTGGAATTTCACTTAACACA 610
Db 886 TCGCCCTTAACCAAGCCGGGAGCTGAGAGAGGATCATGAGCTGCTAAGACATATA 945

```

	Query Match	Similarity	5.8%	Score 180.4	DB 3	Length 6268	
	Best Local	Similarity	52.9%	Pred. NO. 1.9e-37			
	Matches	485	Conservative	0	Mismatches 415	Indels 15	Gaps 4
QY	137	TGCTGATGACACCCAGAGCATTTGAAGTTCACAAAGAGCTCTTGGAAGTGCTGC	196				
Db	469	TGCTTGAAGCCTCGAGATGATAGTGATGAGGAGAAACATGGCCGGGSCCAGGTCGT	528				
QY	197	TGATATGAGTTTGCAACCACTCACTCGTGGAGAGTGACTATTTGGCTCGAGTTTC	256				
Db	529	TTGACCTGGGTCTGTGAACACTCACTCCATAGAGAAAGACTACTTGGCTCGACCTTCT	588				
QY	257	CTGATCACAAAAGATCACGGTGTGGCTGGATCTCTTAAACCATTTGTGAACAGATTA	316				

```

1 RESULT 5
2 US-09-949-016-1488
3 ; Sequence 1488, Application US/09949016
4 ; Patent No. 6812339
5 ; GENERAL INFORMATION:
6 ; APPLICANT: VENTER, J. Craig et al.
7 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
8 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
9 ; FILE REFERENCE: CLO01.07
10 ; CURRENT APPLICATION NUMBER: US/09/949.016
11 ; CURRENT FILING DATE: 2000-04-14
12 ; PRIOR APPLICATION NUMBER: 60/241,755
13 ; PRIOR FILING DATE: 2000-10-20
14 ; PRIOR APPLICATION NUMBER: 60/237,768
15 ; PRIOR FILING DATE: 2000-10-03
16 ; PRIOR APPLICATION NUMBER: 60/231,498
17 ; PRIOR FILING DATE: 2000-09-08
18 ; NUMBER OF SEQ ID NOS: 207012

```

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 1488
LENGTH: 4292
TYPE: DNA
ORGANISM: Human
US-09-949-016-1488

Query Match 5.4%; Score 169.6; DB 3; Length 4292;
Best Local Similarity 51.4%; Pred. No. 1.2e-34;
Matches 472; Conservative 0; Mismatches 434; Indels 12; Gaps 3;

```
OY 127 AAAATCCAGATGCTGATGACACCCAGAGGCACTTTGAAGTTCACAAAGAGCTCTGGG 186
Db 675 AAAGTACCCCTCTTAATGATGACACCGAATACAGCTGACCTGGAGAAACATGCAAGGGA 734
OY 187 AAGGTCTGCTGGATGATGACATTTGCAACCACTCACTGTGGAGAGTACTATTTGGC 246
Db 735 CAAGGTATTATTGACAAAGTGTGAAACCTCAATCTCTGGAGAAAGACTACTTTGGA 794
OY 247 CTCGAGTTTCTGATCACAATAAAGATCAGGATGAGTGGCTGATCTCTAAACCATTTG 306
Db 795 CTTTGTGTTGACAGAAAGCCCTGACGAAATACTGTTGATCTGCTAAAGAAATTAAG 854
OY 307 AAACAGATTGAAGGCCAAAGCACTTGTGTTAAGTTGTGTAATTTCTTCCGCT 366
Db 855 AGACAATGAGAAAC---TTCATGAGCTATTCACTTTAATGTGAAGTTTATCTCTCT 911
OY 367 GACCAACACAACTCCAAAGAACTCACAAGTACTCTTCCGCTGAGGTAAGGAG 426
Db 912 GATCTCTTCAATGACTAAAGATATCACAGATCTTCTGTGCTTCAAGCTCCGGCAG 971
OY 427 GACTTGGCTCAAGGCGAGTTGACGTGTAAATGACACCAAGGCACTCTTGAATTCAC 486
Db 972 GACATTTGCTCTGGCGCGCTGCTGCTCTTTTGTGATCATGATCTCTCTGGATCTCA 1031
OY 487 ATTGTGCAATCTGAGATTGGGGATTTT-----GATGAAGCTTGGACAGAGACTTA 540
Db 1032 ACCCTCAGAGCTGAACTTGGTGAATGACCCAGAAAGAACTGAGCATGCGCTCAGT 1091
OY 541 GCAAAAATAATATACACTCTCAGCAAGCGCACTGAGAGACAAATCGTGGAAATTCAC 600
Db 1092 GAATTCAGTTGGCCCTTACTCAGACTAGAGAGCTGAAAGAGAGTGGCAGAGCTGCAC 1151
OY 601 CATTAACCAATTTGAGCAAAACACGAGAGATCAGATTTCCAGCTCTCTAGAGATTGCC 660
Db 1152 AAAACCCACAGGGGCTTATGCGACAGCAAGTGAATCCAGTTCTTAAGAAATGCAAG 1211
OY 661 CGGCTAGAGATGTATGGAATCCGTTGACCCGCGCAAGAGCAGGAGACAGAAAGATC 720
Db 1212 AGGCTTTCCATGATGATGTGTTGACATCAATGACAGAGACTCAGAAAGTGTGACATC 1271
OY 721 AATTCGCGGTTGCCAAACGCGGAATTCAGTGTTCAGGGTTTCATTAAGATCAATGCG 780
Db 1272 AAGCTGGGCGTGTGCTAATGAGCTTCTCATTTACAAAGACAGACTGCGAATCAATCGT 1331
OY 781 TTCAACTGGGCGCAAGGTGCGGAAGCTGAGCTTCAAGAGAGAGGCTTTCTCATCACTC 840
Db 1332 TTGCTGTGCGCGAAATCTTAAATTTCTTAAGACAGATTACTTCAATTAAGATC 1391
OY 841 CGGCGAGATGCCAATATGTCGTAACAGATTACTTGAATTTCTGTATGCGCAGTGGGAT 900
Db 1392 AGACCGGCGAGAGCTGAAAGTTGAGATGACCATTTCAAACTGCGCAACACCGG 1451
OY 901 TTCTGCAATGCTTTCGGAATAATCTGTGTGAACATGACCTTCTTGAAGCTTTTGA 960
Db 1452 GCAGCGAAAAAGACTATGAAAGTGTGCGAGACATCATCTTTCAAGGCTGTGTT--- 1508
OY 961 GAGCCCAAAACCAAGGCCCAAGCCGTCCTTTAGCCGCGGCTCATCATTTGGTTCAGT 1020
Db 1509 TGTCCAGAGAGCAACCAAGGCTTCTGACCTTGGGCTCAAAATTTGGCTATAGT 1568
OY 1021 GGTGGACTTCAGAGCAG 1038
```

Db 1569 GGCCGACCCACCAAGCAG 1586

RESULT 6

US-09-949-016-208
Sequence 208, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR FILING DATE: 2000-10-20

PRIOR FILING DATE: 2000-10-20

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 208

LENGTH: 4336

TYPE: DNA

ORGANISM: Human

US-09-949-016-208

Query Match 5.4%; Score 169.6; DB 3; Length 4336;
Best Local Similarity 51.4%; Pred. No. 1.2e-34;
Matches 472; Conservative 0; Mismatches 434; Indels 12; Gaps 3;

```
OY 127 AAAATCCAGATGCTGATGACACCCAGAGGCACTTTGAAGTTCACAAAGAGCTCTGGG 186
Db 705 AAAGTACCCCTCTTAATGATGACACCGAATACAGCTGACCTGGAGAAACATGCAAGGGA 764
OY 187 AAGGTCTGCTGGATGATGACATTTGCAACCACTCACTGTGGAGAGTACTATTTGGC 246
Db 765 CAAGGTATTATTGACAAAGTGTGAAACCTCAATCTCTGGAGAAAGACTACTTTGGA 824
OY 247 CTCGAGTTTCTGATCACAATAAAGATCAGGATGAGTGGCTGATCTCTAAACCATTTG 306
Db 825 CTTTGTGTTGACAGAAAGCCCTGAGCAAAAACTGGTTAGATCTCTGTAAGAAATTAAG 884
OY 307 AAACAGATTGAAGGCCCAAGCAAGCTGTTGTTAAGTTGTGTAATTTCTTCCGCT 366
Db 885 AGACAATGAGAAAC---TTCATGAGCTATTCACTTTTAATGTGAAGTTTATCTCTCT 941
OY 367 GACCAACACAACTCCAAAGAAACTCACAAGTACTCTGTCGCGCTGAGGTAAGCAG 426
Db 942 GATCTCTCAATGATGAGTGAAGATATCACAGATTACTTGTGCTTCAAGCTCCGGCAG 1001
OY 427 GACTTGGCTCAAGGAGGTTGACGTGTAAATGACACCAAGGCACTCTTGAATTCACAC 486
Db 1002 GACATTTGCTCTGGCGCGCTGCTGCTCTTTTGTGACTCATGCTCTCTGGAGATCTAC 1061
OY 487 ATTGTGCAATCTGAGATTGGGGATTTT-----GATGAAGCTTGGACAGAGAGCACTTA 540
Db 1062 ACCCTCAGGCTGAATCTTGTGATGATGACCCAGAAAGAAATGAGAGATGACACTCAGT 1121
OY 541 GCAAAAATAATATACACTCTCAGCAAGCACTGAGAGCAAAATGTTGAATTTTAC 600
Db 1122 GAATTCAGTTGGCCCTTACTCAGACTTAAGAGCTGGAAGAGAGTGGCAGAGCTGCAC 1181
OY 601 CATTAACCAATTTGAGCAAAACACGAGAGATCAAGATTTCCAGCTCTCTAGAGATTGCC 660
Db 1182 AAAACCCACAGGGGCTTATGCGCAGACCAAGCTGATTTCCAGTTTGAAGAAATCAAG 1241
OY 661 CGGCTAGAGATGTATGGAATCCGTTGCAACCCGCGCAAGAGCAGGAGAGCTCAAGATC 720
Db 1242 AGCTTTCCATGATGATGTGTTGACTTACATATGCGCAAGAGCTCAGAAAGTGTGACATC 1301
OY 721 AATTCGCGGTTGCCAAACGCGGAATTCAGTGTTCAGGGTTTCATTAAGATCAATGCC 780
```

```
Db 1302 AACGTGGCGCTGTGTCTGAATGAGCTTCTCATTTTAAAGACAGACCTGCAATCATCTGT 1361
Qy 781 TTCAACTGGGCCCAAGGTGCGGAGCTGAGCTTCAAGAGAGAGCGCTTTCATCAACCTC 840
Db 1362 TTTCCTTGGCCGAAAAATCTTAAAAATTTCTATTAAGCAGTAATCTTCAATTAAGATC 1421
Qy 841 CGGCCAGATGCCAAATAGTGGGTACAGAGATACCTTGAATTCCTGATGGCCAGTGGGAT 900
Db 1422 AGACCCGCAAGCTGGAACAGTTGAGATACCAATGGAATTCAAATGCCAAACACCGG 1481
Qy 901 TTTCGCAAGTCTTCTGAAATCTGTGTGAACATGATGCTTCTTAGACTTTTGA 960
Db 1482 GCAGCGAAAAAGACTATGAGAAAGTGTGCTGAGATCACTATTCTTCAAGCTTTGTT--- 1538
Qy 961 GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGTCAATTCCTGGTCACT 1020
Db 1539 TCTCCAGAGCAGCACCAGCAAGTCTTGAATTCCTTGGGGTCCAAATTTCCGTATAGT 1598
Qy 1021 GGTCCGACTCAGAACG 1038
Db 1599 GCGCCGACCCAGACAG 1616
```

RESULT 7
US-09-270-767-1995
Sequence 1995, Application US/09270767

Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270.767
NUMBER OF SEQ ID NOS: 1999-03-17
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1995
LENGTH: 576
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-1995

Query Match 5.3%; Score 165.8; DB 3; Length 576;

Best Local Similarity 61.3%; Pred. No. 4.7e-34; Indels 3; Gaps 1;
Matches 285; Conservative 0; Mismatches 177;

```
Qy 598 CACCATTAACCATTTGGACAACACAGAGATTCAGATTTCCAGCTCTAGAGATTGCC 657
Db 6 CACTTAAGAGATTTGGTCAATCCCGGCGGAGCGGACCTTAACCTTTGAGAGCGCT 65
Qy 658 CGTCGGCTAGAGATGTATGGAATCCGGTTCACCCGCGCAAGCAGGAGGACGAG 717
Db 66 CGAGAGGTGAGCTGTATGCAATGAATCATCCGCGAAGAGATGGAAGGGGTCCG 125
Qy 718 ATCAATCTGGCCGTTGGCAACAGGGAAATCTAGTGTTCAGGGTTTCAATAGATCAAT 777
Db 126 CTTAACTGTGCTGTGGCCACATGGGATCAAGCTTTCAGAAACATCAGCGGATCAAC 185
Qy 778 GCCTTCAACTGGGCCAAGGTGCGGAGCTGAGCTTCAAGAGAGCGCTTTCATCAAG 837
Db 186 ACCTTTCGTCGGGCTAAGATACGAGATTTCTTCAAGGCAAGGATTCCTGGTCAA 245
Qy 838 CTCGGGCGAGATGCCAATAGTGGTACAGAGATACCTTGAATTCCTGATGGCACTGG 897
Db 246 CTCGATCCGAGGAGATGTGATATTACAGAGATACCTGAGATTCCTTCGAGGGTCGC 305
Qy 898 GATTTTCGAGAGCTCTCTGAAAAATCTGTGGAACATCATGCTCTTAGACTTTT 957
Db 306 AACGAGTGCAAAAATCTTGGAAAAATGCGTCAAAAAATACGAGATTTCTTCAGTCACT 365
Qy 958 GAAGAGCCCAA---ACCAAGCCCAAGCCGCTCTTCTTTAGCCGGGGTCAATTTCCG 1014
Db 366 GCGGTACAAAATACGCCAGGCGCAAAAATCGGCTTCTCGGGGGTATGTTATCCCGC 425
```

```
Qy 1015 TTCAAGTGTGCACTCAGAAAGAGTTCCTGACTATGTTAAAGAA 1059
Db 426 TATAGCGGAAAAACCAAGAGAGATTAATGAGTTCGCGAA 470
```

RESULT 8
US-09-270-767-17277

Sequence 17277, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270.767
NUMBER OF SEQ ID NOS: 1999-03-17
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 17277
LENGTH: 576
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-17277

Query Match 5.3%; Score 165.8; DB 3; Length 576;
Best Local Similarity 61.3%; Pred. No. 4.7e-34; Indels 3; Gaps 1;
Matches 285; Conservative 0; Mismatches 177;

```
Qy 598 CACCATTAACCATTTGGACAACACAGAGATTCAGATTTCCAGCTCTAGAGATTGCC 657
Db 6 CACTTAAGAGATTTGGTCAATCCCGGCGGAGCGGACCTTAACCTTTGAGAGCGCT 65
Qy 658 CGTCGGCTAGAGATGTATGGAATCCGGTTCACCCGCGCAAGCAGGAGGACGAG 717
Db 66 CGAGAGGTGAGCTGTATGCAATGAATCATCCGCGAAGAGATGGAAGGGGTCCG 125
Qy 718 ATCAATCTGGCCGTTGGCAACAGGGAAATCTAGTGTTCAGGGTTTCAATAGATCAAT 777
Db 126 CTTAACTGTGCTGTGGCCACATGGGATCAAGCTTTCAGAAACATCAGCGGATCAAC 185
Qy 778 GCCTTCAACTGGGCCAAGGTGCGGAGCTGAGCTTCAAGAGAGCGCTTTCATCAAG 837
Db 186 ACCTTTCGTCGGGCTAAGATACGAGATTTCTTCAAGGCAAGGATTCCTGGTCAA 245
Qy 838 CTCGGGCGAGATGCCAATAGTGGTTCAGAGATACCTTGAATTCCTGATGGCACTGG 897
Db 246 CTCGATCCGAGGAGATGTGATATTACAGAGATACCTGAGATTCCTTCGAGGGTCGC 305
Qy 898 GATTTTCGAGAGCTCTCTGAAAAATCTGTGGAACATCATGCTCTTAGACTTTT 957
Db 306 AACGAGTGCAAAAATCTTGGAAAAATGCGTCAAAAAATCAGAGATTTCTTCAGTCACT 365
Qy 958 GAAGAGCCCAA---ACCAAGCCCAAGCCGCTCTTCTTTAGCCGGGGTCAATTTCCG 1014
Db 366 GCGGTACAAAATACGCCAGGCGCAAAAATCGGCTTCTCGGGGGTATGTTATCCCGC 425
Qy 1015 TTCAAGTGTGCACTCAGAAAGAGTTCGACTATGTTAAAGAA 1059
Db 426 TATAGCGGAAAAACCAAGAGAGATTAATGAGTTCGCGAA 470
```

RESULT 9

US-09-848-294-1
Sequence 1, Application US/09848294
Patent No. 6479640
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas K.
TITLE OF INVENTION: Isolation of A cDNA Encoding A No. 6479640e1
TITLE OF INVENTION: Protein Tyrosine Phosphatase which Localizes to Focal
FILE REFERENCE: CSHL90-04FZA
CURRENT APPLICATION NUMBER: US/09/848,294
CURRENT FILING DATE: 2001-05-03

```
/ PRIOR APPLICATION NUMBER: 09/235,251
/ PRIOR FILING DATE: 1999-01-22
/ PRIOR APPLICATION NUMBER: 08/759,536
/ PRIOR FILING DATE: 1996-12-04
/ PRIOR APPLICATION NUMBER: 08/107,420
/ PRIOR FILING DATE: 1993-08-16
/ PRIOR APPLICATION NUMBER: 07/663,579
/ PRIOR FILING DATE: 1991-03-01
/ PRIOR APPLICATION NUMBER: 07/494,036
/ PRIOR FILING DATE: 1990-03-14
/ NUMBER OF SEQ ID NOS: 13
/ SOFTWARE: FaSTSeq for Windows Version 3.0
/ SEQ ID NO 1
/ LENGTH: 3984
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (24) ... (2765)
US-09-848-294-1
```

Query Match 4.6%; Score 144; DB 3; Length 3984;

Best Local Similarity 50.8%; Pred. No. 8.3e-28; Mismatches 423; Conservative 0; Indels 9; Gaps 3;

```
126 CAAATCCAGATGCTGATGACACCCAGAGGATTTGAAGTTCCAAAGAGCTCTCG 185
116 CAGCATCCACTTTTATGATGGCGTGTACAGACCTTTAAAGTACTAAACAGACCTGG 175
186 GAAGTGTCTGTGATGATGACCTTGAACCACTGCTGAGAGGATGATTTTGG 245
176 CCAGGTTCTTCTGATGATGATGACCAACCTGGGTGATGAGAAATATTTTGG 235
246 CCGGATTTCCGATGACCAAAAGATCCGGTG---TGGCTGATGCTCTAAACCAT 302
236 TTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 295
303 TGTGAACAGATTAGAGGCAAGGACGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 362
296 CAGGAAGAGTTAAAGAGGTTCCCTGTGACCTGATTTTGAAGATTTTAT 355
363 GCTGACCAACCACTCCAGAGAGATCTGATGATGATGATGATGATGATGATGAT 422
356 ACCTGATCCCAACACCTGACAGCAAGCAACCAAGCACTGTATTTTACACTGAA 415
423 GAGGACTTGGCTCAAGGAGGTTGATGATGATGATGATGATGATGATGATGATGAT 482
416 GATGATTTTGGAGAGAGGTTAACTGCTCTTAACTGACAGTGTCTAGCTG 475
483 ACACATTTGCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 539
476 CTATGCGTATCAATCTCATTTTGAAGATTAATTTTCAATCAATCAATCAATCT 535
540 AGCAAAAATTAATCAATCTGACAGCAAGCACTGAGAGCAAAATGTTGAATTTCA 599
536 TTCCGATGCTCTTATATCCGATCAAAATGAGGACTTTTAAACAAAGATGAACTCT 595
600 CCAT---AACCATTTGACAAACACAGAGAAATGATTTCCAGCTCTGAGATTCG 656
596 GATGAGCAGACAGTGGCTAAACATCAAGAGCAATCTGCTATATCAACTAGC 655
657 CCGTGGCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 716
656 GGGGACCTCGACTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 715
717 GATCAATCTGCGCTTCCAAACAGGAAATTTAGTGTTCAGAGTTTCACTAATGCAA 776
716 CCTAAGATTTGAATGCTTCCCGGGGTGTGCTGTGTACGAAATTAATTTGCAAG 775
777 TGGCTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 836
776 TTTCTATCTTGGGTGAATTTCTCAAAATTTTCTTCAAAAGAAAGTTCTTCTACAA 835
```

RESULT 10
US-09-906-779-3

```
/ Sequence 3, Application US/09906779
/ Patent No. 6770466
/ GENERAL INFORMATION:
/ APPLICANT: Shi et al.
/ TITLE OF INVENTION: Human Protein Tyrosine Phosphatase Polynucleotides, Polypeptides,
/ FILE REFERENCE: PTO40P1
/ CURRENT APPLICATION NUMBER: US/09/906,779
/ PRIOR FILING DATE: 2001-07-18
/ PRIOR APPLICATION NUMBER: PCT/US01/01563
/ PRIOR FILING DATE: 2001-01-17
/ PRIOR APPLICATION NUMBER: 60/176,306
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 3
/ LENGTH: 2872
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-906-779-3
```

Query Match 4.6%; Score 143.2; DB 3; Length 2872;

Best Local Similarity 50.3%; Pred. No. 1.2e-27; Mismatches 490; Conservative 0; Indels 16; Gaps 5;

```
83 GACAGAGCGGCCCAACACCTTCAAGAAAATCTGTCTCATCAAAATCCAGATCTGG 142
1 GGCACGAGCGCGCGCGCGCGCCAGAGGCAC-CCTCTACTGCGCGCTTCTCTGCTG 59
143 ATGACACCGAGAGGATTTGAAGTTCCAAAGAGCTCTGGAGAGGCTGCTGATG 202
60 ACGGAGCGAGGTGAGCTGTGACCTGCGAAGCATGCCAAGGCGAGATTTGTTGATC 119
203 CAGTTGCAACCACTCAACCTCGTGAAGGATGATTTTGGCTCGAGTTTCTGATC 262
120 AGATTGTACCACTTGAACCTTGTGAAACAGATTACTTTGGCTTCCAGTTCTCGACT 179
263 ACAAAGATCAAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 322
180 CTGCCAGAGTTGCGCAGCTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 239
323 CAAAGCAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 382
240 GACCTGCTTATGCTTATCACTTTCAGTTAAATCTATTTCTTCAAGAACCAACCTTC 299
383 AAGAGAACTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 442
300 GTGAGAGATTTAAGAGTACCTGTTGTTTGAACATCAGGATGATCAATTTCTGGAA 359
443 GGTGACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 502
360 AATTGAAATGCGCTTATGAAACAGCTGTGAAATGATGATGATGATGATGATGATGATGAT 419
503 TTGGGATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 559
420 TTGGGAGTGGAGCTTCCAGAACACACAGAGCTTGTGTGAGTTTCCGTTCTATTC 479
560 C---TCAGCAAGCAGCTAGAGCAAAATGTTGAATTTTCAACCATTAACCAATTTGAGC 616
480 CAAATCAAGCAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 539
617 AAACACGACGATCAAGATTTCCAGCTCTTGAAGATTTGCCGTGCTAGAGATGATG 676
```


Db	540	AGAGCCCTGCCAGGCGGA	CTTCTTATCTGAATTAAGCGA	AGTGGCTGGAAATG	ATG	599	
Qy	677	GAATCCGGTTGCACCCGGC	CAAGGACAGGAGGACGACGA	ATCAATCTGGCCGT	TGCCA	736	
Db	600	GGGTAGACATC	ACGTTGTCAAGGGAGAGAT	GGCTGGAATATCTCTT	TGGAC	659	
Qy	737	ACACGGGAATTC	TACTGTTTCAAGGTTTCA	CTAAGATCAATGCTTCA	ACTGGGCCAAG	796	
Db	660	CGACAGGCAAT	TATATCTTTGAGGAGCT	TAACAAATATGCGCTTATCTTT	TGGCCTAA	719	
Qy	797	TGCGGAAGCT	GACTTCAAGAGGAAGGC	CTTCTATCAAGTCCGGCCAG	TGCCA	856	
Db	720	TTACCAAAATG	ATTTTAAAGCAAAATG	ACA	CTGTGCTGATGTCAGAGATG	ATG	779
Qy	857	-----G	TCATCCAGGATACCTTGA	ATTCTGATGAGCCAGTGGG	ATTTCTGCA	910	
Db	780	AGGACGTGAG	CAAGACACACGTTTG	TGTTCCGTTAGACAGTGC	CCAGACTTGC	839	
Qy	911	CTTCTGGAAAT	CTGTGTGAACTCA	TGCTTCTTTAGACTTTT	TGAAGCCCAAC	970	
Db	840	ACCTTTGGAAT	GTGACAGTGGACCA	CGGATTTCTTCGACTGGGAC	GCGCAGGAA	899	
Qy	971	CAAAAGCCAA	---GCCGCTCTTT	TAGCCGGGGTCA	TATTTGGTTCA	GTG	1027
Db	900	GCAATATCA	ATAGATCCGACTTT	ATCAGAGCTGGCTCT	CGCTTCA	GTATTCAGTGGCGGA	959
Qy	1028	CTCAGAGC	AGGTT	1041			
Db	960	CAGAAAT	CAAGCT	973			

```

RESULT 11
US-09-949-016-3593
/ Sequence 3593, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949, 016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3593
/ LENGTH: 2156
/ TYPE: DNA
/ ORGANISM: Human
/ US-09-949-016-3593

```

Query Match	4.5%	Score 142.4	DB 3	Length 2156
Best Local	Similarly	Pred. No. 1.6e-27		
Matches 422	Conservative	0	Mismatches 401	Indels 9
				Gaps 3

Qy	126	CAAAATCCAGATGCGGATGACACCCAGAGGCA	TTTGAAGTTCACAAAGAGCTCTGG	185
Db	110	CAGATCCACTTTTATGATGGCGGTGACAGCCTTTAAAGTTACTAAACAAAGACCTGG		169
Qy	186	GAAAGTGTGCTGATGCACTTTGGAACAACCTCAACTCGTGAAGAGTACTATTTGG		245
Db	170	CCAGGTTTTCGTGATATGCTGACAAACCACTGGGTGTACTGAAAAAGATATTTTGG		229
Qy	246	CCTGAGTTTCCTGATCACAAAAGATCAAGGTG--TGGCTGATCTCTTAAACCAT		302
Db	230	TTTACAGCATATGACGACTCCGTGAGACTCTTCCTTAGATGGCTGGAAGCAAGCAAGCAT		289

Qy	303	TCGGAACAAGATTAGAGGCCAAACACAGCTGTTGTTAAAGTTTGAGAAATCTCTCC	3632
Db	290	CAGGAAGCGATTAAAGAGGATTTCCCTGTAACCTGCATTTTGAGTAAGATTTTAT	3498
Qy	363	GCCTGACACACACAACTCCAGAAAGAACTCAAGAGTAACCTGTTGCGTCGACAGTGAA	4222
Db	350	ACCTGATCCCAACACACTGACAGCAAGAAACAACGAGCACTTGTAATTTCTTACAACTGAA	4099
Qy	423	GCAGGACTTGGCTCAAGGACGTTGACGTGTAATGACACGACGACGCTCTTGATTC	4828
Db	410	GATGATATTGCGAAGGAGGTTAACCTGCCCTTTAACTCAACAGGAGTTTCAAGCTC	4658
Qy	483	ACACATTTGCACTGAGATTGGGGATTTTGATGAACCTT---GGACAGAGGACATT	5399
Db	470	CTATGCGCTACATCTCATTTTGGAGACTAATTTCTTCATATCATATCCAGGCTATCT	5289
Qy	540	AGCAAAAAATAAATACATACCTGACAGACGCACTAGAGGACAAAAATCGTGAATTTCA	5999
Db	530	TTCCGATATGACTTTATACCGATCAAAATGAGGACTTTTAACTAAAGTCGATCTCT	5889
Qy	600	CGAT---AACCACTTGGACAAACACGACGAGATCAATTTCCAGCTCCAGAGATTGC	6566
Db	590	GCAATGACGACACAGTGGGCTAAACAAATCGAAGCAAAATCTGCTATATCAACATAGC	6498
Qy	657	CCGTCGGCTAGAGATGATAGAAATCCGGTTGCAACCCGCGCCACAGGACAGGAGGACGAA	7168
Db	650	GCGGACCTTCGACTTCTATGAGATAGAACTGCAACAGTGTAGGATTCGCAATTTAGA	7099
Qy	717	GATCAATCTGCGCGTTGCCACACGCGAAATTCATGTTTCAAGGCTTCACTAAGATCAA	7768
Db	710	CCTAATGAATTGAAATTCCTTCGCGCGGTTGTCGTGTACGAAATAATTTGCAACAG	7658
Qy	777	TGCTTCACTGAGGCGCAAGGTGCGGAGCTGAGCTTCAAGAGAAAGGCTTTCATCAA	8368
Db	770	TTTCTATCTCTGGGTGAACATTTCTAAAAATTTCTTTCAAAAGGAAAGTTCTTATACA	8289
Qy	837	GCTCGGCGCAGATGCCAATAGTGCGTACACAGATACCTTGAATTCCTGATGCGCAGTCG	8966
Db	830	TCAGCGACAGAAACAGGCTGAATTCACAGGAAACATTTGTGCTTCAACATGCTGAATTA	8889
Qy	897	GGATTTCTGCAAGTCTCTCGAAAAATCTGTGTTGAACATATAGCTCTTT	948
Db	890	CGGATTTGCAAAACTGTGAAAAATCTGTGTTGAGCACATAGCTTCTTT	941

RESULT 12
 US-09-799-451-201
 Sequence 201, Application US/09799451
 Patent No. 6783969
 GENERAL INFORMATION:
 APPLICANT: Tang, Y. Tom
 APPLICANT: Zhou, Ping
 APPLICANT: Goodrich, Kyle
 APPLICANT: Asundi, Vinod
 APPLICANT: Ren, Feiyun
 APPLICANT: Zhang, Jie
 APPLICANT: Xue, Aidong J.
 APPLICANT: Zhao, Qing A.
 APPLICANT: Wang, Jians-Rui
 APPLICANT: Ma, Yundong
 APPLICANT: Yamazaki, Victoria
 APPLICANT: Chen, Rui-hong
 APPLICANT: Wang, Zhiwei
 APPLICANT: Wang, Dunrui
 APPLICANT: Yang, Yonghong
 APPLICANT: Wehrman, Tom
 APPLICANT: Ghosh, Reena
 APPLICANT: Dimaac, Radoje T.
 TITLE OF INVENTION: Polypeptides
 TITLE OF INVENTION: No. 6783969e1 Nucleic Acids and
 FILE REFERENCE: 803
 CURRENT APPLICATION NUMBER: US/09/799,451
 CURRENT FILING DATE: 2001-03-05


```
Db      199  CCATCCGAAAGCAATACGGCTCTGCAAGTCTGTTGGCATGTGGCGGCAACGTAGACTG 258
Qy      2773 GTGGACTTCAGCATCCGACGTGAGAAATCAATTGTCTGGAACCTGCTAGAGAAATTCAAA 2832
Db      259  GGGGATCACCCTGATATCCGCCGAGACCAAGTTGCCGTTATCTGCTGCGAAAGTTCAAG 318
Qy      2833 AACAGCAACGGGTGGGAGAAAGTGTGGGTGTTGACAAACTTCGCCCTTCTTCTAC 2892
Db      319  AATAGCTCGGCTGGCAAAAGCTCTGGGTGTGTTACGTCCTTTGTCTGTACTTTTAC 378
Qy      2893 AATACACACGAGCAATCATCCCTTGCCAGCTGCTGCTGCGGCTACTGCT 2948
Db      379  AAAAGCTACGAGACGAGTTGCACTGGCTAGTCTGCCGCTATTGGGTTACAGGT 434
```

RESULT 15

```
US-09-621-976-15728
; Sequence 15728, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15728
; LENGTH: 184
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15728
```

```
Query Match      3.4%; Score 105.2; DB 3; Length 184;
Best Local Similarity 99.1%; Pred. No. 4.4e-18;
Matches 105; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1692 GTGGTTCAAGACAGTGAAGCAAGAGGACGCCCATGCCGGAAGCACTGAAAAGTCTCAT 1751
Db      79  GTGGTTCAAGACAGTGAAGCAAGAGGACGCCCATGCCGGAAGCACTGAAAAGTCTCAT 138
Qy      1752 ATTCCGAAATTTGAACCTTTGCACAAAATTTCATTAATTTTCTC 1797
Db      139 ATTCCGAAATTTGADCTTTGCACAAAATTTCATTAATTTTCTC 184
```

Search completed: May 22, 2006, 09:18:20
Job time : 543.44 sec

This Page Blank (uspto)

181 CTTGGGAAGTGCTGTGATGCAATTTGCAACCACTCAACTCTGTGAAAGTGACTAT 240
| | | | |
Db 229 CTTGGGAAGTGCTGTGATGCAATTTGCAACCACTCAACTCTGTGAAAGTGACTAT 288
| | | | |
Qy 241 TTTGGCTCGAGTTTCTGTATCAAAAAGATCAAGTGTGCTGTGATCTCTAAAACC 300
| | | | |
Db 289 TTTGGCTCGAGTTTCTGTATCAAAAAGATCAAGTGTGCTGTGATCTCTAAAACC 348
| | | | |
Qy 301 ATTGTGAAACAGTTTGAAGGCCAAAGCAAGTGTGTTAAGTTTGTGTAATCTTT 360
| | | | |
Db 349 ATTGTGAAACAGTTTGAAGGCCAAAGCAAGTGTGTTAAGTTTGTGTAATCTTT 408
| | | | |
Qy 361 CGGCTGACCAACACCACTCCAAAGAACTCAAGGTAAGTGTGCTGTGACGAGTG 420
| | | | |
Db 409 CGGCTGACCAACACCACTCCAAAGAACTCAAGGTAAGTGTGCTGTGACGAGTG 468
| | | | |
Qy 421 AAGCAGACTTGTGCTCAAGGCAAGTGTGACGTGTATGACACAGCGCAGCTCTTGAT 480
| | | | |
Db 469 AAGCAGACTTGTGCTCAAGGCAAGTGTGACGTGTATGACACAGCGCAGCTCTTGAT 528
| | | | |
Qy 481 TCACATTTGTGCAATCTGAGATTTGGGGAATTTGAGGAAGCTTTGACAGAGCACTTA 540
| | | | |
Db 529 TCACATTTGTGCAATCTGAGATTTGGGGAATTTGAGGAAGCTTTGACAGAGCACTTA 588
| | | | |
Qy 541 GCATAAATTAATACATACCTCAGCAAGACGCACTAGAGGCAAAAATCGTGAAATTCAC 600
| | | | |
Db 589 GCATAAATTAATACATACCTCAGCAAGACGCACTAGAGGCAAAAATCGTGAAATTCAC 648
| | | | |
Qy 601 CATTAACCACTTGGACAAACACCAAGAAATCAGATTTCCAGTCTTGAAGTTGCCGT 660
| | | | |
Db 649 CATTAACCACTTGGACAAACACCAAGAAATCAGATTTCCAGTCTTGAAGTTGCCGT 708
| | | | |
Qy 661 CGGCTGAGATGATGGAATCCGGTTGCAACCGGCAAGGACAGGGAAGGCAAGAGATC 720
| | | | |
Db 709 CGGCTGAGATGATGGAATCCGGTTGCAACCGGCAAGGACAGGGAAGGCAAGAGATC 768
| | | | |
Qy 721 AATCTGCGCTTGGCAACAGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCGC 780
| | | | |
Db 769 AATCTGCGCTTGGCAACAGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCGC 828
| | | | |
Qy 781 TTCACCTGGGCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGCTTTCTCATCAAGCTC 840
| | | | |
Db 829 TTCACCTGGGCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGCTTTCTCATCAAGCTC 888
| | | | |
Qy 841 CGGCGAGATGCCAATGTGCTACAGGAATCTTGAATCTCTGATGCGAGTCCGGAT 900
| | | | |
Db 889 CGGCGAGATGCCAATGTGCTACAGGAATCTTGAATCTCTGATGCGAGTCCGGAT 948
| | | | |
Qy 901 TTCTGCAAGTCTTCTGGAATCTGTGTTGAACATCATGCTCTTTAGACTTTTGA 960
| | | | |
Db 949 TTCTGCAAGTCTTCTGGAATCTGTGTTGAACATCATGCTCTTTAGACTTTTGA 1008
| | | | |
Qy 961 GAGCCCAACCAAGCCCAAGCCCTCTTTAGCCGGGGGTCAATCTTGGTTCAAGT 1020
| | | | |
Db 1009 GAGCCCAACCAAGCCCAAGCCCTCTTTAGCCGGGGGTCAATCTTGGTTCAAGT 1068
| | | | |
Qy 1021 GGTGGAATCTCAGAAAGGTTCTGACTATGTTAAAGAAAGAGACATTAAGAGGTGACG 1080
| | | | |
Db 1069 GGTGGAATCTCAGAAAGGTTCTGACTATGTTAAAGAAAGAGACATTAAGAGGTGACG 1128
| | | | |
Qy 1081 TTTGAAAGAGACAGACAGATTTCAATTATCCGAGCCTTGCTTACAGCTTACAGAA 1140
| | | | |
Db 1129 TTTGAAAGAGACAGACAGATTTCAATTATCCGAGCCTTGCTTACAGCTTACAGAA 1188
| | | | |
Qy 1141 CTGAATTCGGAAGTGTGAGCAGTCTCAGCAGACACCAAGCTTACATTTGGAAGGT 1200
| | | | |
Db 1189 CTGAATTCGGAAGTGTGAGCAGTCTCAGCAGACACCAAGCTTACATTTGGAAGGT 1248
| | | | |
Qy 1201 GCGGAATTCACAGGGGGCCAGAGCTGCGGCGAGAAAGGAACGAAGTTTCCGCGGG 1260
| | | | |
Db 1249 GCGGAATTCACAGGGGGCCAGAGCTGCGGCGAGAAAGGAACGAAGTTTCCGCGGG 1308
| | | | |
Qy 1261 GAGCCGAGGTGCAACCCGAGCCTTGCGCGAGAGAAAGCCCGCGGGTTAAACAAGCAGGCG 1320
| | | | |

1309 GAGCCGAGGTGCAACCCGAGCCTTCGCGAGAGAAAGCCCCGGGGTTAAACAAGAGGCG 1368
| | | | |
Qy 1321 GACGAGCGCCTCGGCGCCCAACGAGAAAGAGAGAGTGTGTTAAGATAGAACCCG 1380
| | | | |
Db 1369 GACGAGCGCCTCGGCGCCCAACGAGAAAGAGAGAGTGTGTTAAGATAGAACCCG 1428
| | | | |
Qy 1381 CAGAGTAAACCTCAGCCCCCGAGCAAGCAGACAGGCTCCCTGACTGCGAGTCTCACTT 1440
| | | | |
Db 1429 CAGAGTAAACCTCAGCCCCCGAGCAAGCAGACAGGCTCCCTGACTGCGAGTCTCACTT 1488
| | | | |
Qy 1441 TCGAGCTGTCTGTGAATCTCGAGGGGGAGTGGCCCTTCGCAACGTGACTTTGTCTCC 1500
| | | | |
Db 1489 TCGAGCTGTCTGTGAATCTCGAGGGGGAGTGGCCCTTCGCAACGTGACTTTGTCTCC 1548
| | | | |
Qy 1501 AACCTGAGCCCCGAGACCAAGAGGCTCTCCCTTGATACGCCCGCTGTGAATGACAG 1560
| | | | |
Db 1549 AACCTGAGCCCCGAGACCAAGAGGCTCTCCCTTGATACGCCCGCTGTGAATGACAG 1608
| | | | |
Qy 1561 GCTGCCCCCGAGACGAGATGAGGATGAGGCGGAGAGAAAGATTCCCACTGATATA 1620
| | | | |
Db 1609 GCTGCCCCCGAGACGAGATGAGGATGAGGCGGAGAGAAAGATTCCCACTGATATA 1668
| | | | |
Qy 1621 GCGTACTTCAATAGCTAAGGAAGTGTCTACACCGAGCGAACATATCTGAAGATCTCGAA 1680
| | | | |
Db 1669 GCGTACTTCAATAGCTAAGGAAGTGTCTACACCGAGCGAACATATCTGAAGATCTCGAA 1728
| | | | |
Qy 1681 GTTATCACTTGTGTTTCAAGACACATGAGCAAAAGGAAGCCATGCCGGAACACAGC 1740
| | | | |
Db 1729 GTTATCACTTGTGTTTCAAGACACATGAGCAAAAGGAAGCCATGCCGGAACACAGC 1788
| | | | |
Qy 1741 AAAAGTCAATATTCGCAATTTTGAACCTTTCGCAAAATTCATACTAATTTTCTCAAG 1800
| | | | |
Db 1789 AAAAGTCAATATTCGCAATTTTGAACCTTTCGCAAAATTCATACTAATTTTCTCAAG 1848
| | | | |
Qy 1801 GAAATTTGACAAAGACTTGCCTTGGGAAAGGCCCTCAATGCCCCAAATCAAGATTAC 1860
| | | | |
Db 1849 GAAATTTGACAAAGACTTGCCTTGGGAAAGGCCCTCAATGCCCCAAATCAAGATTAC 1908
| | | | |
Qy 1861 CAAGAATCCGGAGATGATGTAAGAAACATTCAGGGCATGAACACTGCGCGCTCAC 1920
| | | | |
Db 1909 CAAGAATCCGGAGATGATGTAAGAAACATTCAGGGCATGAACACTGCGCGCTCAC 1968
| | | | |
Qy 1921 CTGTGGAAGCAGAGAGGCTTGGAGGCCCTTGAGAAATGGAATCAAGAGCTCCGGCGG 1980
| | | | |
Db 1969 CTGTGGAAGCAGAGAGGCTTGGAGGCCCTTGAGAAATGGAATCAAGAGCTCCGGCGG 2028
| | | | |
Qy 1981 CTGTGGAAGCAGAGAGCTTGAAGCTGCAAGAAAGTGTGTTACCTTACCGCTCAACACC 2040
| | | | |
Db 2029 CTGTGGAAGCAGAGAGCTTGAAGCTGCAAGAAAGTGTGTTACCTTACCGCTCAACACC 2088
| | | | |
Qy 2041 TTCTCTCGCGGCACTGCAACCGGCTCATGCACTTCAAGAGAGTCTGTGAGGGCGTGC 2100
| | | | |
Db 2089 TTCTCTCGCGGCACTGCAACCGGCTCATGCACTTCAAGAGAGTCTGTGAGGGCGTGC 2148
| | | | |
Qy 2101 AAACACACCCGCGAGCCGACGACTTCAGGGGACTGCGGAGCCGCTTTGGCAGAGATC 2160
| | | | |
Db 2149 AAACACACCCGCGAGCCGACGACTTCAGGGGACTGCGGAGCCGCTTTGGCAGAGATC 2208
| | | | |
Qy 2161 ACGAGATGTGTGCAACAGTCTCAAGGTACATGATGAGAGAAATTTCCAGAGCTG 2220
| | | | |
Db 2209 ACGAGATGTGTGCAACAGTCTCAAGGTACATGATGAGAGAAATTTCCAGAGCTG 2268
| | | | |
Qy 2221 CAGGAATCTCAAGAAATTTGATTGGCAATGCAATCTTGTGGTTCCGGGGAAGGAGTTC 2280
| | | | |
Db 2269 CAGGAATCTCAAGAAATTTGATTGGCAATGCAATCTTGTGGTTCCGGGGAAGGAGTTC 2328
| | | | |
Qy 2281 ATCCGTCTGGGAGGCTCAGCAAGCTCTCGGGGAAGGGGCTTCAGACGCGCATGTTCTTC 2340
| | | | |
Db 2329 ATCCGTCTGGGAGGCTCAGCAAGCTCTCGGGGAAGGGGCTTCAGACGCGCATGTTCTTC 2388
| | | | |
Qy 2341 CTGTTCAACAGAGTCTGTGTTATCAGAGCCGGGGGCTGACGGCTCCAAATCAAGTTTAA 2400
| | | | |

2389 CTGTTCAAGACGCTCTGCTATATACAGAGCCGGGGCTGACGGCTCTCAATCAGTTTAA 2448
QY 2401 GTCCAGGGGACAGTCCCGCTCTATATGCAATGAGAGCGAAGCGAGTGGGG 2460
DB 2449 GTCCAGGGGACAGTCCCGCTCTATATGCAATGAGAGCGAAGCGAGTGGGG 2508
QY 2461 GTGCCCCACTGCTGACCCCTCCGGGGCAGAGCGAGTCCATCATCTGTCGCGCAGATTCT 2520
DB 2509 GTGCCCCACTGCTGACCCCTCCGGGGCAGAGCGAGTCCATCATCTGTCGCGCAGATTCT 2568
QY 2521 CGGTCGAGATGAGAGAGTGGGTTGAGAGCATTCAGATGCCATTGACCTGGCGAGAG 2580
DB 2569 CGGTCGAGATGAGAGAGTGGGTTGAGAGCATTCAGATGCCATTGACCTGGCGAGAG 2628
QY 2581 AGCAGAGCCCGCCCTCTGAGTTCTTGCCAGAGCCCTCTCAACAAGTCCCTGAT 2640
DB 2629 AGCAGAGCCCGCCCTCTGAGTTCTTGCCAGAGCCCTCTCAACAAGTCCCTGAT 2688
QY 2641 GAAGCCAGCCGCTGACAGAGATGACCTGAGCCCTCGGCAATGCTG 2700
DB 2689 GAAGCCAGCCGCTGACAGAGATGACCTGAGCCCTCGGCAATGCTG 2748
QY 2701 GAGCGCAGAGCCCGCAGAGCCGCGCAACACATGCTGCACTGTCGCGCAGCAGC 2760
DB 2749 GAGCGCAGAGCCCGCAGAGCCGCGCAACACATGCTGCACTGTCGCGCAGCAGC 2808
QY 2761 AGGTCCTCATGTTGATCTTACAGCATTCGAGTGGAGATTCAGTTGCTGGAACCTGCTG 2820
DB 2809 AGGTCCTCATGTTGATCTTACAGCATTCGAGTGGAGATTCAGTTGCTGGAACCTGCTG 2868
QY 2821 AGGAAATTCAAAACAGCAACGGGTGGAGAGCTGTTGCTGCAAACTCTGC 2880
DB 2869 AGGAAATTCAAAACAGCAACGGGTGGAGAGCTGTTGCTGCAAACTCTGC 2928
QY 2881 CTGTTCTTCTACAAATCACAACGAGCAATCATCCCTTTCGAGCCCTCTGCTGCGC 2940
DB 2929 CTGTTCTTCTACAAATCACAACGAGCAATCATCCCTTTCGAGCCCTCTGCTGCGC 2988
QY 2941 TACTGCTCAGCATCCCTCTGAGTCCGAGAACATTCAGAAAGATTCAGCTGTTCAAGCTG 3000
DB 2989 TACTGCTCAGCATCCCTCTGAGTCCGAGAACATTCAGAAAGATTCAGCTGTTCAAGCTG 3048
QY 3001 CACTTCAAGTCCCACTACTACTTCAAGGGCGAAGCGAGTACAGTTCGAAAGTGG 3060
DB 3049 CACTTCAAGTCCCACTACTACTTCAAGGGCGAAGCGAGTACAGTTCGAAAGTGG 3108
QY 3061 ATGGAAGTATCGCAGTGCACAGCTGCTGCGAGCCCAAGTGTGAGCCACAA 3120
DB 3109 ATGGAAGTATCGCAGTGCACAGCTGCTGCGAGCCCAAGTGTGAGCCACAA 3168
QY 3121 GAGTCTCTGTGTAT 3135
DB 3169 GAGTCTCTGTGTAT 3183

RESULT 2
US-10-921-707-18
Sequence 18, Application US/10921707
Publication No. US20050003447A1
GENERAL INFORMATION:
APPLICANT: INCYTE PHARMACEUTICALS, INC.
APPLICANT: LAL, Preeti
APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: BANDMAN, Olga
APPLICANT: CORLEY, Neil C.
APPLICANT: GUEGLER, Karl J.
APPLICANT: PATTERSON, Chandra
APPLICANT: AZIMZAI, Yalda
APPLICANT: BAUGHN, Mariah R.
TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
FILE REFERENCE: PF-0594 PCT

CURRENT APPLICATION NUMBER: US/10/921,707
CURRENT FILING DATE: 2004-08-19
PRIOR APPLICATION NUMBER: US/09/786,797
PRIOR FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: 09/156,470; unassigned; 60/131,321
PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PERL Program
SEQ ID NO 18
LENGTH: 4687
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 1320252
US-10-921-707-18

Query Match 99.9%; Score 3131.8; DB 9; Length 4687;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3133; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGGAAATATGAGCAGAGCCGACCCAGATTCAGACTGGGGCCCCGAAATTTG 60
DB 90 ATGGGAAATATGAGCAGAGCCGACCCAGATTCAGACTGGGGCCCCGAAATTTG 149
QY 61 GGGATCAGTACCTGGGAAGTGGACAGAGCCGCCCAACCTTCAGGAAATCTG 120
DB 150 GGGATCAGTACCTGGGAAGTGGACAGAGCCGCCCAACCTTCAGGAAATCTG 209
QY 121 TCCATCAAAATCAGATGCTGATGACACCCAGAGGCAATTTGAAGTTCCAGAAAGCT 180
DB 210 TCCATCAAAATCAGATGCTGATGACACCCAGAGGCAATTTGAAGTTCCAGAAAGCT 269
QY 181 CCGGGAAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 270 CCGGGAAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 329
QY 241 TTTGGCTGAGTTTCTGATCAAAAAGATCAGGTTGGCTGATCTCTAAACC 300
DB 330 TTTGGCTGAGTTTCTGATCAAAAAGATCAGGTTGGCTGATCTCTAAACC 389
QY 301 ATTGTGAAACAGATTGAAAGCCAAAGCACTGTTGTTAAAGTTGTGTAATCTTT 360
DB 390 ATTGTGAAACAGATTGAAAGCCAAAGCACTGTTGTTAAAGTTGTGTAATCTTT 449
QY 361 CCGCTGACCAACAACAATCCAGAGAACTCAAGATACCTGTTGCGCTGCAAGT 420
DB 450 CCGCTGACCAACAACAATCCAGAGAACTCAAGATACCTGTTGCGCTGCAAGT 509
QY 421 AGCAGGACTTGGCTCAAGGAGGTTGACGTGTAATGACACAGAGGCTCTCTGATT 480
DB 510 AGCAGGACTTGGCTCAAGGAGGTTGACGTGTAATGACACAGAGGCTCTCTGATT 569
QY 481 TCACACATTTGCAATCTGAGATTGGGGATTTTGAATGAAGCTTGGACAGAGCACTTA 540
DB 570 TCACACATTTGCAATCTGAGATTGGGGATTTTGAATGAAGCTTGGACAGAGCACTTA 629
QY 541 GCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
DB 630 GCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 689
QY 601 CATTAACCATTTGGAACAAACACAGAGAAATGAGATTCAGCTCTAGAGATTGCGGT 660
DB 690 CATTAACCATTTGGAACAAACACAGAGAAATGAGATTCAGCTCTAGAGATTGCGGT 749
QY 661 CGGCTAGAGATGATGAATCCGTTGCAACCCGCGCAAGAGCAGGAGATC 720
DB 750 CGGCTAGAGATGATGAATCCGTTGCAACCCGCGCAAGAGCAGGAGATC 809
QY 721 AATCTGGCGTTGCCAACAACGGGAATTTCAAGTTTCAAGGTTTCACTAAGTCAATGCC 780
DB 810 AATCTGGCGTTGCCAACAACGGGAATTTCAAGTTTCAAGGTTTCACTAAGTCAATGCC 869

840 781 TTCAACTGGGCCAAGGTGCGAAGCTGAGCTTCAAGAGAGCGCTTCTCATCAAGCTC
870 TTCAATTTGGGCCAAGGTGCGAAGCTGAGCTTCAAGAGAGCGCTTCTCATCAAGCTC
900 841 CGGCCGATGCGCAATAGTGGTACAGAGTACCTTGAATTCCTGATGCGCAGTGGGAT
930 CGGCCGATGCGCAATAGTGGTACAGAGTACCTTGAATTCCTGATGCGCAGTGGGAT
960 901 TTCTGGAATCTCTTGGAAAAATCTGTGTGAACATCATGCGCTTCTTGAACCTTTTGA
990 TTCTGGAATCTCTTGGAAAAATCTGTGTGAACATCATGCGCTTCTTGAACCTTTTGA
1020 961 GAGCCCAACCAAGGCCCAAGCCGCTCTCTTACCGGGGGTCAATTCGCTGAGT
1050 GAGCCCAACCAAGGCCCAAGCCGCTCTCTTACCGGGGGTCAATTCGCTGAGT
1080 1021 GGTGGACTCAGAGAGAGTCTCGACTATGTTAAAGAGAGACATAGAGGTGAG
1110 GGTGGACTCAGAGAGAGTCTCGACTATGTTAAAGAGAGACATAGAGGTGAG
1140 1081 TTTGAAAGAGAGAGAGATTCATTCGCGAGCGCTTCAAGCGCTACAGAA
1170 TTTGAAAGAGAGAGAGATTCATTCGCGAGCGCTTCAAGCGCTACAGAA
1200 1141 CTGAATTCGAAAGTGTGAGAGAGTCTCAGAGAGAGAGAGAGAGTCTTGAAGAGAGT
1230 CTGAATTCGAAAGTGTGAGAGAGTCTCAGAGAGAGAGAGAGAGTCTTGAAGAGAGT
1260 1201 GCCGAATTCAGAGAGAGAGAGTGTGCGAGAGAGAGAGAGAGAGTGTGCGAG
1290 GCCGAATTCAGAGAGAGAGAGTGTGCGAGAGAGAGAGAGAGAGTGTGCGAG
1320 1261 GAGCCGGGGTGGACCGGAGCGCTGCGAGAGAGAGAGAGAGAGCGGGTAAAGAGAGAG
1350 GAGCCGGGGTGGACCGGAGCGCTGCGAGAGAGAGAGAGAGCGGGTAAAGAGAGAG
1380 1321 GACGAGAGCGGCTCGAGCGGCTCAGAGAGAGAGAGAGAGAGTGTGAAGATAGAGAG
1410 GACGAGAGCGGCTCGAGCGGCTCAGAGAGAGAGAGAGAGAGTGTGAAGATAGAGAG
1440 1381 CAGAGTAACTTACGAGCGGCTCGAGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAG
1470 CAGAGTAACTTACGAGCGGCTCGAGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAG
1500 1441 TCCGAGCTGTGTGAACTCGAGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAG
1530 TCCGAGCTGTGTGAACTCGAGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAG
1560 1501 AACCTGAGCGCGAGAGAGAGAGAGTGTGAGAGAGAGAGAGAGTGTGAGAGAGAG
1590 AACCTGAGCGCGAGAGAGAGAGAGTGTGAGAGAGAGAGAGAGTGTGAGAGAGAG
1620 1561 GCGTGGCGCGAGAGAGAGAGAGTGTGAGAGAGAGAGAGAGTGTGAGAGAGAG
1650 GCGTGGCGCGAGAGAGAGAGAGTGTGAGAGAGAGAGAGAGTGTGAGAGAGAG
1680 1621 GGGTACTTATGCTAAGAGAGTGTACACCGAGAGAGAGAGTGTGAGAGAGTGTGAG
1710 GGGTACTTATGCTAAGAGAGTGTACACCGAGAGAGAGAGTGTGAGAGAGTGTGAG
1740 1681 GTTATCACTTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
1770 GTTATCACTTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
1800 1741 AAAAGTCTATATTTCCGAAATTTGAACCTTTGACAAATTTCTATATTTCTCAAG
1830 AAAAGTCTATATTTCCGAAATTTGAACCTTTGACAAATTTCTATATTTCTCAAG
1860 1801 GAAATGAG
1890 GAAATGAG
1920 1861 CAAAGATCGGCGAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

2009 1950 CAAAGATCGGCGAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
1980 1921 CTGTGAG
2069 2010 CTGTGAG
2040 1981 CTGAG
2129 2070 CTGAG
2100 2041 TTCTCTGCGGCACTGACCGGCTCATGCACTAACAGAGAGAGAGAGAGAGAGAG
2189 2130 TTCTCTGCGGCACTGACCGGCTCATGCACTAACAGAGAGAGAGAGAGAGAGAG
2160 2101 AAAACACACCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
2249 2190 AAAACACACCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
2220 2161 ACGAGATGAG
2309 2250 ACGAGATGAG
2280 2221 CACGAATCAAG
2369 2310 CACGAATCAAG
2340 2281 ATCCGCTGAG
2429 2370 ATCCGCTGAG
2400 2341 CTGTTCAAG
2489 2430 CTGTTCAAG
2460 2401 GTCCAGGAG
2549 2490 GTCCAGGAG
2520 2461 GTGCCAG
2609 2550 GTGCCAG
2580 2521 CGGTCCGAG
2669 2610 CGGTCCGAG
2640 2581 AGCAG
2729 2670 AGCAG
2700 2641 GAGGAG
2789 2730 GAGGAG
2849 2790 GAGGAG
2820 2761 AGGCTCTCATGAG
2909 2850 AGGCTCTCATGAG
2880 2821 AGGAATTCAG
2969 2910 AGGAATTCAG
2940 2881 CTGTTCTTCAAG
3029 2970 CTGTTCTTCAAG
3000 2941 TACTGCTCATGAG

Accession	Sequence	Position
Db	3030 TACTGCGTCACACATCCCTCTGAGTCCGAAACATCCGAAAAGACTACGTGTTCAAGTGG	3089
Qy	3001 CACTTCAAGTCCACGCTCTACTACTTCAGGCGGAAAGCAGTACACGTTTGAAAGGTGG	3060
Db	3090 CACTTCAAGTCCACGCTCTACTACTTCAGGCGGAAAGCAGTACACGTTTGAAAGGTGG	3149
Qy	3061 ATGGAAGTGATCCGCAGTGCCACCAACGCTCTGCTCGCGACCCACGTTTGAGCCACAA	3120
Db	3150 ATGGAAGTGATCCGCAAGTCCACCAACGCTCTGCTCGCGACCCACGTTTGAGCCACAA	3209
Qy	3121 GAGTCTCTGTGTAT	3135
Db	3210 GAGTCTCTGTGTAT	3224

```

RESULT 3
US-10-450-763-564
/ Sequence 564, Application US/10450763
/ Publication No. US20050196754A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc
/ TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
/ FILE REFERENCE: 790CIP3/US
/ CURRENT APPLICATION NUMBER: US/10/450,763
/ CURRENT FILING DATE: 2003-06-11
/ PRIOR APPLICATION NUMBER: PCT/US01/08631
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: 09/540,217
/ PRIOR FILING DATE: 2000-03-31
/ PRIOR APPLICATION NUMBER: 09/449,167
/ PRIOR FILING DATE: 2000-08-23
/ NUMBER OF SEQ ID NOS: 60736
/ SOFTWARE: Custom
/ SEQ ID NO 564
/ LENGTH: 3187
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SIMILAR
/ LOCATION: (49)..(1371)
/ OTHER INFORMATION: 92% homologous to Homo sapiens CDEP, accession number
/ US-10-450-763-564
/ OTHER INFORMATION: AB008450, Smith-Waterman Score=2090.

```

Query Match	99.5%	Score 319.2	DB 10	Length 3187
Best Local Similarity	99.9%	Pred. No. 0		
Matches 3132	Conservative 0	Mismatches 3	Indels 1	Gaps 1
QY	1	ATGGGAGAAAATAGACAGAGGCCCA	CCCCCAGATTCACGACTCTGGGGCCCCCGGAAAATTCG	60
Db	49	ATGGGAGAAAATAGACAGAGGCCCA	CCCCCAGATTCACGACTCTGGGGCCCCCGGAAAATTCG	108
QY	61	GGGATCTAGTACTTGGAACTGTGGA	CAGAAAGCCGCCCAACACTTCACGAAAATCTCGTG	120
Db	109	GGGATCTAGTACTTGGAACTGTGGA	CAGAAAGCCGCCCAACACTTCACGAAAATCTCGTG	168
QY	121	TCCATCAAAATCCAGATGCTGGAAT	GACACCCAGGAGGACTTTGAAGTTCCACAAAGACT	180
Db	169	TCCATCAAAATCCAGATGCTGGAAT	GACACCCAGGAGGACTTTGAAGTTCCACAAAGACT	228
QY	181	CCTGGGAAGGCTGCTGGATGTCAGT	TTCACCAACCTCAACTCTGTTGGAAGTGACAT	240
Db	229	CCTGGGAAGGCTGCTGGATGTCAGT	TTCACCAACCTCAACTCTGTTGGAAGTGACAT	288
QY	241	TTTGGCCTCGAGTTCTCTGATCAC	AAAAAGATCAAGGTGTGGCTGGATCTCTTAAACC	300
Db	289	TTTGGCCTCGAGTTCTCTGATCAC	AAAAAGATCAAGGTGTGGCTGGATCTCTTAAACC	348
QY	301	ATTGTGAACAGATTGAAAGGCCAA	AGCAGTGTGTTGAAGTTGTGTGAAATTCCTT	360
Db	349	ATTGTGAACAGATTGAAAGGCCAA	AGCAGTGTGTTGAAGTTGTGTGAAATTCCTT	408
QY	361	CCGCTTGACACACACAACTCCAA	AGAGACTCACAGATGCTGTTCGCGCTGCAGGTG	420

Db	409	CCGCTGACCAACAACAACCTCCAGAAAGAACTCACAAAGTACCTGTTCGCGTGCAGGTC	468
OY	421	AAGCAGACTTGGCTCAAGCAGGTGACGTGTATGACACACGCGACGCTCTCTTGATT	480
Db	469	AAGCAGACTTGGCTCAAGCAGGTGACGTGTATGACACACGCGACGCTCTCTTGATT	528
OY	481	TCACACATTTGTGCACATCTGAGATTGGGATTTTGTATGAAGCTTGGACAGAGACCTTA	540
Db	529	TCACACATTTGTGCACATCTGAGATTGGGATTTTGTATGAAGCTTGGACAGAGACCTTA	588
OY	541	GCAAAAATAAATCATCTCCACAGAGACGCACTAGAGGCAAAATGCTGAATTTAC	600
Db	589	GCAAAAATAAATCATCTCCACAGAGACGCACTAGAGGCAAAATGCTGAATTTAC	648
OY	601	CATAACACATTTGACAAACACACAGAGAACTCAGATTTCCAGCTCTAGAAATTTCCGCT	660
Db	649	CATAACACATTTGACAAACACACAGAGAACTCAGATTTCCAGCTCTAGAAATTTCCGCT	708
OY	661	CGGCTAGAGATGTATGAAATCCGGTTGACCCCGCCAAAGACAGGGAAGGACGAAAGTC	720
Db	709	CGGCTAGAGATGTATGAAATCCGGTTGACCCCGCCAAAGACAGGGAAGGACGAAAGTC	768
OY	721	AATCTGCGCCTTGCACAACGCGGAATTTCTATGTTTTCAGGGTTTCACTTAAGATCAATCC	780
Db	769	AATCTGCGCCTTGCACAACGCGGAATTTCTATGTTTTCAGGGTTTCACTTAAGATCAATCC	828
OY	781	TTCAACTGGGCGCAAGGTCCGAGAGCTGAGCTTCAAGAGAGAGCGCTTCTCATCAAGTCC	840
Db	829	TTCAACTGGGCGCAAGGTCCGAGAGCTGAGCTTCAAGAGAGAGCGCTTCTCATCAAGTCC	888
OY	841	CGGCGAATGTCAAATAAGTCGCTACACAGGATACCTTGGAATTCCTGATGCGCAGTCGGAT	900
Db	889	CGGCGAATGTCCAATAAGTCGCTACACAGGATACCTTGGAATTCCTGATGCGCAGTCGGAT	948
OY	901	TTCTGCAAGTCTCTTGGAATCTGTGTGAACATCATGCTTCTTTAGACTTTTGA	960
Db	949	TTCTGCAAGTCTCTTGGAATCTGTGTGAACATCATGCTTCTTTAGACTTTTGA	1008
OY	961	GAGCCCAAACCAAAGCCCAAGCCCGCTCTTTAGCCGGGGGTACATCTTCGGTTCAGT	1028
Db	1009	GAGCCCAAACCAAAGCCCAAGCCCGCTCTTTAGCCGGGGGTACATCTTCGGTTCAGT	1068
OY	1021	GGTGGAATCTGAAGACAGGTTCTCGACTATGTAAAGAGAGGACATTAAGAGGTGAC	1080
Db	1069	GGTGGAATCTGAAGACAGGTTCTCGACTATGTAAAGAGAGGACATTAAGAGGTGAC	1128
OY	1081	TTTGAAGAGAGACACAGCAGATTCATCTATCCGAGCCTTTCACAGCCTACAGAA	1140
Db	1129	TTTGAAGAGAGACACAGCAGATTCATCTATCCGAGCCTTTCACAGCCTACAGAA	1188
OY	1141	CTGAATTCGGAAGGTCTGAGCAGTCTCAGAGAGACACAGCCTTACATTTGGAGAGAG	1199
Db	1189	CTGAATTCGGAAGGTCTGAGCAGTCTCAGAGAGACACAGCCTTACATTTGGAGAGAG	1248
OY	1200	TGCCGAATCTTCCAGGGGGCCAGAGCTGCGCGAGAGAAAGAACCGAAGTTTTCGCGCG	1255
Db	1249	TGCCGAATCTTCCAGGGGGCCAGAGCTGCGCGAGAGAAAGAACCGAAGTTTTCGCGCG	1308
OY	1260	GGAGCCGGGGTTCGCACTCCGAGCCTTGCGCAGAGAGAAAGCCCGCGGGTTAACAGCAGGC	1319
Db	1309	GGAGCCGGGGTTCGCACTCCGAGCCTTGCGCAGAGAGAAAGCCCGCGGGTTAACAGCAGGC	1368
OY	1320	GGAAGGAGCGCCTCGGCGGCCACGAGAGAAAGAGAGAGGTCTGTAAAGATTAGAACCA	1379
Db	1369	GGAAGGAGCGCCTCGGCGGCCACGAGAGAAAGAGAGAGGTCTGTAAAGATTAGAACCA	1428
OY	1380	GCAGAGTAACTTCAGGCCCCCGCAGCCAGACAGCAGGCTCCCTGACTGGCAGTCTCTCACT	1439
Db	1429	GCAGAGTAACTTCAGGCCCCCGCAGCCAGACAGCAGGCTCCCTGACTGGCAGTCTCTCACT	1488
OY	1440	TTCCGAGCTGCTGTGAATCTGCAGGGGGAGGTGGCCCTTGCACAGTGAATCTTGTCTCC	1499

1489 TTCCGAGCTGTCTGTGTAATCCGACGGGGAGTGGCCCTCGCCAACTGATCCTTGTCTCC 1548
1500 CAACCTGACCCCCGACACCAAGACGGGCTCTCCCTTGATCAGCCGGCTGTGAATGACA 1559
1549 CAACCTGACCCCCGACACCAAGACGGGCTCTCCCTTGATCAGCCGGCTGTGAATGACA 1608
1560 GGCCTGCCCCCGGACGAGATGAGATGAGGCGCGGAGAGAGATTTCCCACTGATTA 1619
1609 GGCCTGCCCCCGGACGAGATGAGATGAGGCGCGGAGAGAGATTTCCCACTGATTA 1668
1620 ACGCTACTTCACTAAGAGAGTGTCTACACGAGCCGACATATCTGAGAGATCTGCA 1679
1669 ACGCTACTTCACTAAGAGAGTGTCTACACGAGCCGACATATCTGAGAGATCTGCA 1728
1680 ACGTATCACTGCTGTGTTCAGAGCAGTGTGACAAAGAGAGAGCCGATGCGGAGACCT 1739
1729 ACGTATCACTGCTGTGTTCAGAGCAGTGTGACAAAGAGAGAGCCGATGCGGAGACCT 1788
1740 GAAAGTCTCATATTCGCAATTTTGAACCTTTGCAAAATTTTCATATTAATTTTCTCA 1799
1789 GAAAGTCTCATATTCGCAATTTTGAACCTTTGCAAAATTTTCATATTAATTTTCTCA 1848
1800 GGAATTTGACACCACTTCCCTGTGGAGAGCCGCTCAATGCCCAATACAGATTA 1859
1849 GGAATTTGACACCACTTCCCTGTGGAGAGCCGCTCAATGCCCAATACAGATTA 1908
1860 CCAAGAAATCGGAGATGTCTGTGAAGACATTCAGGGGATGAAACCTGGCGGCTCA 1919
1909 CCAAGAAATCGGAGATGTCTGTGAAGACATTCAGGGGATGAAACCTGGCGGCTCA 1968
1920 CCGTGTGAGACACAGGAGAGCCCTTGGAGGCCCTGAGAAATGAAATGAAGCTCCGAGG 1979
1969 CCGTGTGAGACACAGGAGAGCCCTTGGAGGCCCTGAGAAATGAAATGAAGCTCCGAGG 2028
1980 GCTGTGAGAACTTCTGTAGAGACTTTGAGCTGTGCAAGAGGTGTGTACTTACCGTCAAC 2039
2029 GCTGTGAGAACTTCTGTAGAGACTTTGAGCTGTGCAAGAGGTGTGTACTTACCGTCAAC 2088
2040 CTTCCCTCTGCGGCGCACTGCAACCGGCTCATGCACTCAAGACAGGTCTGTGAGCGGCTG 2099
2089 CTTCCCTCTGCGGCGCACTGCAACCGGCTCATGCACTCAAGACAGGTCTGTGAGCGGCTG 2148
2100 CAAACACCAACCCGCGGACGACGCGCACTTCAGGAGCTGCCGAGCGCTTTGGCAGAGAT 2159
2149 CAAACACCAACCCGCGGACGACGCGCACTTCAGGAGCTGCCGAGCGCTTTGGCAGAGAT 2208
2160 CACGAGATGTGTGACCACTCCACGCTACAGATCAAGATGAGAAATTTCCAGAGCT 2219
2209 CACGAGATGTGTGACCACTCCACGCTACAGATCAAGATGAGAAATTTCCAGAGCT 2268
2220 GCACGAACTCAAGAAAGATTTGATTTGGCAATCTTGTGTTCGGGAGAGGAGATT 2279
2269 GCACGAACTCAAGAAAGATTTGATTTGGCAATCTTGTGTTCGGGAGAGGAGATT 2328
2280 CATCCGTCTGGGAGGCTTCAGCAAGCTCTCGGAGAAAGGAGCTTCAGCAGCGCATGTTCT 2339
2329 CATCCGTCTGGGAGGCTTCAGCAAGCTCTCGGAGAAAGGAGCTTCAGCAGCGCATGTTCT 2388
2340 CCGTGTTCACAGAGTCTGTCTATACAGAGCCGGGGGCTGACGGCTCCCAATCACTTTAA 2399
2389 CCGTGTTCACAGAGTCTGTCTATACAGAGCCGGGGGCTGACGGCTCCCAATCACTTTAA 2448
2400 AGTCCACGGGAGAGCTCCGCTCTATAGGCAATGACATTTGAGAGAGAGAGAGAGAGAGG 2459
2449 AGTCCACGGGAGAGCTCCGCTCTATAGGCAATGACATTTGAGAGAGAGAGAGAGAGG 2508
2460 GGTGCCCCCATGCTGCTGACCTTCGGGAGGACGCGCATTCATCATCTGTGGCGCCAGTTT 2519
2509 GGTGCCCCCATGCTGCTGACCTTCGGGAGGACGCGCATTCATCATCTGTGGCGCCAGTTT 2568
2520 TCGGTCCGAGATGAGAAATGTGGTTGAGACATCCAGATGCGCATTTGACCTGGCGGAGAA 2579
2569 TCGGTCCGAGATGAGAAATGTGGTTGAGACATCCAGATGCGCATTTGACCTGGCGGAGAA 2628

2580 GAGCAGACAGCCCCGCGCTGAGTTCCTGGCAGACAGCCCCCTGACAAAGTCCCCCTGA 2639
2629 GAGCAGACAGCCCCGCGCTGAGTTCCTGGCAGACAGCCCCCTGACAAAGTCCCCCTGA 2688
2640 TGAAGCCACCGGCGGTGACACAGAGTCAAGAGATGACCTGAGCGCTCGGCAATGCTCT 2699
2689 TGAAGCCACCGGCGGTGACACAGAGTCAAGAGATGACCTGAGCGCTCGGCAATGCTCT 2748
2700 GGAAGGCCAGAGCCCGGACCCGGGCAACATGTGTGACGCTGTGGCACCGGAAAC 2759
2749 GGAAGGCCAGAGCCCGGACCCGGGCAACATGTGTGACGCTGTGGCACCGGAAAC 2808
2760 CAGCGTCTCAGTGTGACCTTCAGATGCACTGTGAGATCAGTTGTCTGAAACCTGCT 2819
2809 CAGCGTCTCAGTGTGACCTTCAGATGCACTGTGAGATCAGTTGTCTGAAACCTGCT 2868
2820 GAGAAATTTCAAAAACAGCAACGGGTGCAAGAGCTGTGGTGTGACAACTTCTG 2879
2869 GAGAAATTTCAAAAACAGCAACGGGTGCAAGAGCTGTGGTGTGACAACTTCTG 2928
2880 CCGTGTCTCTCAAAATCAACAGAGCAATCATCCCTTGCAGCTGCTGTGG 2939
2929 CCGTGTCTCTCAAAATCAACAGAGCAATCATCCCTTGCAGCTGCTGTGG 2988
2940 CTACTGCTCACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTACGTGTTCAGCT 2999
2989 CTACTGCTCACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTACGTGTTCAGCT 3048
3000 GCACCTTCAAGTCCACAGTCTTACTTCAAGGCGGAGAGAGAGTACAGTTGAAAGTGT 3059
3049 GCACCTTCAAGTCCACAGTCTTACTTCAAGGCGGAGAGAGAGTACAGTTGAAAGTGT 3108
3060 GATGAGAGATCCGAGAGGCCAGCAGCTGTGCTGCGACCCCACTGTGAGCCACA 3119
3109 GATGAGAGATCCGAGAGGCCAGCAGCTGTGCTGCGACCCCACTGTGAGCCACA 3168
3120 AGAGTCTCTTGTGTAT 3135
3169 AGAGTCTCTTGTGTAT 3184

RESULT 4
US-09-925-297-220
; Sequence 220, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 3094
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (4)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-220

Query Match 45.7%; Score 1431.2; DB 3; Length 3094;
Best Local Similarity 99.8%; Pred. No. 0;

Matches 1441; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1692 GTGGTTTCAAGCACTAGTGAAGCAAGGAGCGCCATGCGGAGACTGAAAAGTCTCAT 1751
DB 62 GTGGTTTCAAGCACTAGTGAAGCAAGGAGCGCCATGCGGAGACTGAAAAGTCTCAT 121
QY 1752 ATTCCCGAATTTTGAACCTTTGCACAATTTTCACTAAATTTTCTCAAGGAATTTGACA 1811
DB 122 ATTCCCGAATTTTGAACCTTTGCACAATTTTCACTAAATTTTCTCAAGGAATTTGACA 181
QY 1812 ACGACTTCCCTGTGGAGAGCGCGCTCAATGCCCAATCAAGATTACCAAGAAATCGG 1871
DB 182 ACGACTTCCCTGTGGAGAGCGCGCTCAATGCCCAATCAAGATTACCAAGAAATCGG 241
QY 1872 CGATGTATGTGAAAGCAATTCAGGGATGAAACACTGCGGCTCACTGTGAAAGCA 1931
DB 242 CGATGTATGTGAAAGCAATTCAGGGATGAAACACTGCGGCTCACTGTGAAAGCA 301
QY 1932 CAGGAGGCGCTTGGAGGCGCTGAGAAATGAAATCAAGAGCTCCGCGGCGCTGAGAACTT 1991
DB 302 CAGGAGGCGCTTGGAGGCGCTGAGAAATGAAATCAAGAGCTCCGCGGCGCTGAGAACTT 361
QY 1992 CTGCAAGACTTTGAGCTGCAAGAGGTGTGTTACTACCGCTCAACACTTCTCTGCG 2051
DB 362 CTGCAAGACTTTGAGCTGCAAGAGGTGTGTTACTACCGCTCAACACTTCTCTGCG 421
QY 2052 GCCATGCAACCGGCTCATGCACTCAAGCAGGTCTTGAGCGGCTGTGCAACCAACCC 2111
DB 422 GCCATGCAACCGGCTCATGCACTCAAGCAGGTCTTGAGCGGCTGTGCAACCAACCC 481
QY 2112 GCCGAGCGCGCGCACTTCAAGGAGCTGCGAGCGGCTTTGGCAGAGATCAAGAGATGT 2171
DB 482 GCCGAGCGCGCGCACTTCAAGGAGCTGCGAGCGGCTTTGGCAGAGATCAAGAGATGT 541
QY 2172 GGCAAGCTTCAAGCTGATGATCAAGATGAGAAATTTTCAAGCTGCAAGAACTCAA 2231
DB 542 GGCAAGCTTCAAGCTGATGATCAAGATGAGAAATTTTCAAGCTGCAAGAACTCAA 601
QY 2232 GAAAGATTTGATTTGCAATCTTTGCTTCCGGAAGGGAAGTTCACTCGTCTGGG 2291
DB 602 GAAAGATTTGATTTGCAATCTTTGCTTCCGGAAGGGAAGTTCACTCGTCTGGG 661
QY 2292 CAGGCTCAGCAAGCTCTCGGGGAAAGGGGCTCAGCAGCGGATGTTCTTCCGTTCAAGCA 2351
DB 662 CAGGCTCAGCAAGCTCTCGGGGAAAGGGGCTCAGCAGCGGATGTTCTTCCGTTCAAGCA 721
QY 2352 CGTCTGCTATTAAGCAAGCGGGGCTGAGCGGCTTCAATCAGTTTAAAGTCCAAGGCA 2411
DB 722 CGTCTGCTATTAAGCAAGCGGGGCTGAGCGGCTTCAATCAGTTTAAAGTCCAAGGCA 781
QY 2412 GCTCCGCTCTATGCAATGATGAGAGGGAAGCAAGTGGGGGTTGCCCACTG 2471
DB 782 GCTCCGCTCTATGCAATGATGAGAGGGAAGCAAGTGGGGGTTGCCCACTG 841
QY 2472 CTTGACCTTCCGGGCGCAGGCGGCTCATCATCTGTCGCGGCACTTCTGCTCCGAGAT 2531
DB 842 CTTGACCTTCCGGGCGCAGGCGGCTCATCATCTGTCGCGGCACTTCTGCTCCGAGAT 901
QY 2532 GGAGAAATGAGTTGAGAGCAATCCAGATGAGCACTTGAAGTGAAGAGCAAGCAAGCC 2591
DB 902 GGAGAAATGAGTTGAGAGCAATCCAGATGAGCACTTGAAGTGAAGAGCAAGCAAGCC 961
QY 2592 CGGCGCTGATTTCTTGCCGAGCGGCGGCTTGAACAAGTCCCTGATGAGCAAGCC 2651
DB 962 CGGCGCTGATTTCTTGCCGAGCGGCGGCTTGAACAAGTCCCTGATGAGCAAGCC 1021
QY 2652 GGCTGACCAAGATCAGAGATGAGCTGAGCGCTGCGCAATCGCTGAGCGCGCAAGG 2711
DB 1022 GGCTGACCAAGATCAGAGATGAGCTGAGCGCTGCGCAATCGCTGAGCGCGCAAGG 1080
QY 2712 CCGGCAACCGGCGCAACAAATGTGCAAGTGTGCTGCGCAACCGCAACCGCTTTCAT 2771
DB 1081 CCGGCAACCGGCGCAACAAATGTGCAAGTGTGCTGCGCAACCGCAACCGCTTTCAT 1140

QY 2772 GTGGAATTGAGAGTGCAGTGAAGATCAGTTGTCTGSAACCTGCTGAGAAATTTCAA 2831
DB 1141 GTGGAATTGAGAGTGCAGTGAAGATCAGTTGTCTGSAACCTGCTGAGAAATTTCAA 1200
QY 2832 AAAACAGCAAGGATGAGAGAGCTGTGGGTGTGTTTCAAACTTCTGCTTCTTCA 2891
DB 1201 AAAACAGCAAGGATGAGAGAGCTGTGGGTGTGTTTCAAACTTCTGCTTCTTCA 1260
QY 2892 CAATCAACAGAGCAATCATCCCTTGGCAAGCTGCTTCTGCTGCTTCTGCTGCTAC 2951
DB 1261 CAATCAACAGAGCAATCATCCCTTGGCAAGCTGCTTCTGCTGCTTCTGCTGCTAC 1320
QY 2952 CATCCCTTGAAGTCCGAGAAATCCAGAAAGCTACGTTTCAAGCTGCACTTCAAGTC 3011
DB 1321 CATCCCTTGAAGTCCGAGAAATCCAGAAAGCTACGTTTCAAGCTGCACTTCAAGTC 1380
QY 3012 CCAAGTCTACTACTTCAAGGCGGAGAGCAAGTACAGTTTCAAAAGGTGATGAAATGAT 3071
DB 1381 CCAAGTCTACTACTTCAAGGCGGAGAGCAAGTACAGTTTCAAAAGGTGATGAAATGAT 1440
QY 3072 CCGAGTGCACCAAGCTCTGCTTGGGAGCCCAAGTGTGAGCCCAAGAGTCTTGT 3131
DB 1441 CCGAGTGCACCAAGCTCTGCTTGGGAGCCCAAGTGTGAGCCCAAGAGTCTTGT 1500
QY 3132 GTAT 3135
DB 1501 GTAT 1504

RESULT 5
US-10-172-118-1536
; Sequence 1536, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yundong
; APPLICANT: Linley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1536
; LENGTH: 3997
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 014808
; DATABASE ENTRY DATE: 2001-06-18
; US-10-172-118-1536

Query Match 33.1%; Score 1037.4; DB 7; Length 3997;
Best Local Similarity 60.2%; Pred. No. 6.7e-302;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

QY 21 GCCGACCCCAAGATCAAGACTGAGGAGCGCGGAAAATTCGGGATCATGACTTGGAAAG 80
DB 150 GCAGACTGACGAGATCGCTTGGGTCCAGACCCCTGTGGAGTTAGCACCTTGAAGCC 209
QY 81 TGACAGAAAGCGCGCCCAACACTTCAAGAAATCTCGTCCATCAAAATCAGATGCT 140
DB 210 TGAGCAAGACTCTTGGCCCAAGATGCAAGAGAGCACTGCACTCAAGATTAAGTGT 269
QY 141 GGATGACACCAAGAGGCAATTTGAAATTCACAAAGAGTCTTGGAGAGGTGCTGCTGA 200

270 GGACAAACACATGAAATATTTGACATGAGCCTTAATGCGATGCGCAGATATTAATGAC 329
201 TCGAGTTTGCAACCACTCAACCTGCGTGAAGGTGACTATTTTGGCTCGAGTTTCTGA 260
330 ACAAGTGTGAGACGCTTTAACTGCTGAGATGATGCTACTTCTGAGATGAGATTTTAA 389
261 TCACAAAAGATCACGCGTGTGCTGATCTCTAAACCCTATTTGTAACAGATTTAGAG 320
390 TACTGAGCTCACTGATTTGGCTTGAACCTATGAACCATCATTTAGCAATATGAGAG 449
321 GCCAAAGACGCTTGTGTTAAGTTTGTGTAATCTTTCCGCTGACACACAACT 380
450 GCCAAAGATGTGTGCTTGCCTGACTGTAATTTTTCACCTGATCTGTGTAGCT 509
381 CCAGAAGAACTCACAGGATCTGTGCGCTGACAGTGAAGCAGAGACTTGGCTCAAG 440
510 ACAAGAAATATACAGATCTTGTGCTTTCCTTGAACCTTAAGAGAGACTGCTGAA 569
441 CAGTTGACGTGTATGACACGACGCTCTTGTGATTTTCAACATTTGTGCAATCTGA 500
570 GGGTTTGACCTGTGCTGACACGACGCGCTTCTCACTCCATCTCTGCACTCGGA 629
501 GATTGGGATTTTGTGATGAAGCTTTGACAGAGACCTTAGCAAAAATTAATATAC 560
630 AATAGAGATTAAGATGAACGCTGACGAGACCTCAAGTGAAGAGATATTTGCT 689
561 TCAGCAAGACGACCTAGAGACAAATGCTGAATTTTCAACATTAACCATTTGGA 620
690 TGGCCAGCAGCACTGCTGAGAAATGATCTAGATTTTCACTAGAGACAGTGGCCAG 749
621 ACCAGCAGATCAGATTTTCACTGCTCTAGAGATTTGCTGCTGCTGAGATGATG 680
750 ACTGCTGATGCGAATTTTCAAGGTGCTCGAAATTTGCTCGAAATGTAAGCGCAT 809
681 CCGGTTGCAACCGGCAAGAGACGAGAGACAGAGATCAATCTGCGCTTGCCTCAAC 740
810 CAGATTTTCACTGCTGCTGACAGAGAGACCAAGATTTCACTGCTGCTTCCACAT 869
741 GGGATTTCTAGATTTTCAAGGCTTCTCAATGATCAATGCTTCACTGAGGCTGAG 800
870 GGGTGTACTGTGTTTCAAGGACCAACCAATCACTTCACTGCTGCTGAGTCCG 929
801 GAGCTGACCTTCAAGAGAGAGGCTTCTCATCAAGCTTCCGCTGAGATGCAATGTC 860
930 TAAACTTAAGCTTCAAGAGAGAGATTTCTTATCAATCTTCACTGAGGCTTCAAG 989
861 GATCCAGATTAAGTGAATTTCTGATGAGGCTGAGTCTGAGATTTCTGAGATCTT 920
990 TTAACGAGACATTAAGAAATTTTGTGAGTGAAGATGAAGATTTGAAGATTTG 1049
921 AATCTGTGTTGAACATGCTTCTTATGACTTTTGAAGGCTTCAAGGCTTCAAGGCT 980
1050 GATTTGTGTGAGATACACACTTTTATGACTTTTGAACCACTTAAGGCTTCAAGGCT 1109
981 GCCGCTCTTCTTGAAGGCTGATCAATTTGCTGATGAGTCTGAGTCTGAGAGAG 1040
1110 AGCCGCTTCTTCAAGGCTGCTCTCTTCAAGATCAAGTGAAGACTCAAGAACT 1169
1041 TCTGATCTGTTGAAGAGAGAGACATTAAGAGTGAAGTGAAGAGAGAGAGAGAG 1100
1170 AGTAGATTTTCAAGAGAGAGAGATTAAGAGATTTCAATGAAGAGAGAGAGAGAG 1229
1101 GATTCATTTATCCGAGCTTCTTCAAGCTTCAAGAGCTTGAAGTGAAGTGAAGTGA 1160
1230 GACCCACAGTCT 1271
1161 GAGATCTCAAG 1220
1272 ACAGAGATCTCAATCT 1331
1221 GAGCTGCGGAG 1280
1332 CTTTAACTGCT 1391

1281 CCTGCGCGAG 1340
1392 CAGCAGCT 1451
1341 CAGCAG 1397
1452 CAGTGAAG 1511
1398 CCGGAGCGCAG 1457
1512 ACTCAGGCTGTGCTCAGGCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1571
1458 CTGCAAG 1517
1572 GAGCCCTGAGTCTGAG 1622
1518 CAGCAG 1577
1623 ACAGGCTCTATCT 1682
1578 CAGTGAAG 1637
1683 CTGCAAG 1739
1638 GGAAGTGTCTTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1697
1740 AGAGATTTCTGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1799
1698 TCAGAGCAG 1757
1800 CCGCAG 1859
1758 GAATTTTGAACCTTGTGCAAAATTTTCAATTAATTTTCTCAAGAAATTTGAGCAAG 1817
1860 CAACATGATCT 1919
1818 TGCCTGTGAG 1874
1920 GGCAGCTCTGAG 1979
1875 TGTATGCTGAG 1934
1980 CATCTGCTCAG 2039
1935 CGAGGCTTGTGAG 1994
2040 CAGGCTCTTACAG 2099
1995 CAGAGATTTGAGCTGAG 2054
2100 CAGAGATTTGAGCTGAG 2159
2055 ACTGACCGGCTCAG 2114
2160 CATCAG 2219
2115 GAGCCAGCGGCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2174
2220 CGGAGACATGAG 2279
2175 ACAGCTTCAAGGATGAG 2234
2280 CACACTACAGCAATTTCTATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2339
2235 AGATTTGATTTGAG 2294
2340 GAGCTGTGTGAG 2399
2295 CCTCAGCAAGCTCTGAG 2354
2400 CTTTCAAG 2459

2355 CCTGCTATACAGACCGCGGGGCTGACGGCCTCCATCACTTTAAATTCACGGGACGT 2414
2460 GTTGTCTATACAAAGAAAGAGTTGACGGGACCAAGCACTTCCGGATCCGGGGCTCTCT 2519
2415 CCGGCTCTATAGCATGCAATTGAGAGAGCGAAAGCAAGTGGGGGTGCCCCCTGCTCT 2474
2520 TCCCTCTCAAGGCACTGCTGGTGGAAAGTGAATACAGAGTGGCTCTTCCACACTGTTT 2579
2475 GACCCCTCGGGGGCCAGGGGAGTCCATCATCTGTCGGCCGCAAGTTCTGCTCCAGATGA 2534
2580 CACATCTACCGGCTAGAAACAAATCTGTGTGGCCAGCACTGGCTGGAGAAAGA 2639
2535 GAAGTGGGTGAGGACATCCAGATGCGCATTTGACTGGCGGAGAGAGACAGCCCGC 2594
2640 GAAGTGGATGCTGGAACCTGAATCTCCGGATCCAGAGCCAAAGAGTGGCGGTGACGGC 2699
2595 CCTGTGATCTTGGCCAGCA-----GCCCTCTGACAAAGTCCCTGATGAAGCAC 2648
2700 CCTGTGACTGCGAGCGGCACTGTGTGCACTGCTCCCGCAGATCCCGCAAGAGTATC 2759
2649 CGGGGCTGACCAAGATCAAGATGACCTGAGGGCCTCCGCGACATCGCTGAGGGCCA 2708
2760 TCTG---GAGCAGAGATCAAGATGATGCTGGGGGTGTCGCAAGCTCCCTGAGGGGCA 2816
2709 GCGCCGCAACCGGCGCAACAATGATGTCAGTGTGTGGCAACGCAACACAGCGTCTC 2768
2817 TGGCCAGACACGGGGCCAAACACCAATGCACTGTGTGTGGTACCGGAACACAGCGTGT 2876
2769 CATGTGAGACTTCAAGCATCGCAGTGAAGAAATCAAGTTGTGTGAAACCTGTGAGAAAT 2828
2877 CAGGGCAGACCAAGTGCACGTGTGAGAAACAGCTTTCAAGATATCTGTGAAGAAAGTT 2936
2829 CAAAAACAGAAAGGGTGGAGAAAGCTGTGGGTGTGTTCACAACTTCTGCTGTTCTT 2888
2937 CAAAAACAGATGCTGTGGAGAAAGCTGTGGGTGTGTTCACAACTTCTGCTGTTCTT 2996
2889 CTCAATATCACACAGGACCAATATCCCTTGGCAGCGCTGCTGTGCTCGGCTACTGCT 2948
2997 CTACAAACCTCATCAGATACATACCACTGGCCAGCTCCCGCTGTGGGCTACAGCT 3056
2949 CACCATCCCTCTGAGTCCGAGAAATCAAGAAAGTACGTGTTCAGCTGCACTTCAA 3008
3057 GAGCATCCCGAGGAGGCGATGATACAAAGAACTATGTTTCAAGCTCCAGTTCAA 3116
3009 GTCCCACTCTTACTTACTTCAAGGGCGGAAACGATACACTTGGAAAGGTGATGAAGT 3068
3117 ATCCCACTCTTACTTCTTCCGGGCTGTGAGCAAGTACATTTGAAAGGTGATGAGGT 3176
3069 GATCCGACAGTGCACCAAGCTGTGCTGTGCGAGCCC 3103
3177 GATCCAGGGGGCCAGCAAGCTCAAGCCGGAGGGGCC 3211

RESULT 6
US-10-342-887-1536
Sequence 1536, Application US/10342887
Publication No. US20040058340A1
GENERAL INFORMATION:
APPLICANT: Dai, Hongyue
APPLICANT: He, Yudong
APPLICANT: Lineley, Peter S.
APPLICANT: Mao, Mao
APPLICANT: Roberts, Christopher J.
APPLICANT: Van 't Veer, Laura Johanna
APPLICANT: Van de Vijver, Marc J.
APPLICANT: Bernardes, Rene
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-188-999
CURRENT APPLICATION NUMBER: US/10/342,887
PRIOR FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/298,918
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/380,710

PRIOR FILING DATE: 2002-05-14
PRIOR APPLICATION NUMBER: 10/172,118
PRIOR FILING DATE: 2002-06-14
NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 1536
LENGTH: 3997
TYPE: DNA
ORGANISM: Homo sapiens
US-10-342-887-1536

Query Match 33.1%; Score 1037.4; DB 8; Length 3997;
Best Local Similarity 60.2%; Pred. No. 6,7e-302;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

21 GCGGACCCCGAGATCAAGACTGGGGGCCCCGGGAAATTCGGGGATCATGTAATTTGAAACG 80
150 GCAAGCTGACGGGATGCGCTTGGGTGCCAGACCCCTGTGGAGTTAGACCTTTGAGCC 209
81 TGGACAGAAAGCCGCCCAACACCTTCAAGAAATCTGTGCTCATCAAAATCCAGATGCT 140
210 TGGGACAGATCTTGTGCCAGAAATGCAAGAAAGCACCTGACCTCAGAGTAAAGTGTCT 269
141 GAATGACACCCAGAGAGCATTTGAAGTTCACAAAGACTCTGGAGAGGTGCTGCTGA 200
270 GGAACAACCATGGAATATTTGACATTTGAGCTTAATGCGATGGCCAGGATTTACTGAC 329
201 TGCAGTTTGAACCACTCACTCGTGGAGGTGATATTTTGGCTCGAGTTTCTCTGA 260
330 ACAAGTGTGAAACGTTTAAACCTGTGAGAAATGACTACTTGGGATGAGATTTCAAAA 389
261 TCACAAAAGATACAGGTGTGCTGATCTCTAAACCATTTGTGAAACAGATTGAAG 320
390 TACTCATGCTCTATGATTTGGCTTGAACCTATGAAACCATATTAAGCAATACGAG 449
321 GCCAAAGCAGTGTGTTAAAGTTGTGTAATTTCTTCCGCTGACCAACACAACT 380
450 GCCAAAGAAATGTGTGCTTGCCTGACCTGATGTAATTTTTCACCTGATCTGCTGAGCT 509
381 CCAAGAAACTCAAGATCTGTTCGCGCTGACAGTGAAGACAGACTTGGCTCAAG 440
510 ACAAGAAAGATTAACAAGATTAATCTGTTGCTTGGCAACTTAAAGAGACCTGTGAAAG 569
441 CAGTTGACGTGTAAGACACACAGCGAGCTCTTGATTTCAACATTTGCAATCTGA 500
570 GCGTTGACCTGTGCTGACACACAGCGGCTTCTCAAGTCCACTCTCGTAGTCGGA 629
501 GATTTGGGATTTTGAAGACCTTGAACAGAGACCTTAGCAAAAATTAATATACATACC 560
630 AATAGAGATTAAGATGAAGCGCTGAGACGAGAGACCTCAAGATGAACGAGATTTGCC 689
561 TCAGCAAGACGACTGAGAGCAAAATCGTGAATTTCACTATACCACTTTGACAAAC 620
690 TGGCCAGACGACTGCTTGAAGAAATTAAGATTTCAAGAAACGCTGGGCGCAGAC 749
621 ACCAGAGATCAGATTTCCAGCTCTTAGAGATTTGCCGCTGAGTGAATGATGAAT 680
750 ACTGCTGAGTGGGATTTCCAGGTGCTCAAAATTTGCTCAAAATTTGAAATTTGACGAT 809
681 CCGGTTGCAACCGGCGCAAGAGCAAGGAGGACAGAGATCAATCTGCGGTTGCCAACAC 740
810 CAGATTTCAATGCTTCTGACAGGAGAGAACCAAGATTTCAACTGCGAGTTTCCACAT 869
741 GGGAAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCACTGAGGCGCAAGTGC 800
870 GGGTGTACTGTGTTCAGGGCAACCAAAATCAACATTTCACTGTCGCAAGGTGC 929
801 GAAGTGAAGCTCAAGAGAGGCGCTTCTCATCAAGCTCCGGCGAGATGCCAATAGTGC 860
930 TAACTAAGCTTCAAGAGAAAGATTTCTTATCAAACTTCAATCCAGAGTTCAATGACC 989
861 GTACAGAGATACCTTGAATTCCTGATGCGCAAGTGGGATTTCTGCAAGTCTTCTGAA 920
990 TTACAGAGACACTTGAATTTTGTGGGTAGTAGATGAATGAAGAACTTCTGAA 1049

Db 3177 GATCCAGGGGGCCAGCAGCTCAGCCGGAGGGGCC 3211

RESULT 7
US-10-802-432-16
; Sequence 16, Application US/10802432
; Publication No. US20040185489A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann-La Roche Inc.
; TITLE OF INVENTION: Transcriptional Activity Assay
; FILE REFERENCE: 21574
; CURRENT APPLICATION NUMBER: US/10/802,432
; CURRENT FILING DATE: 2004-03-17
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 3997
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: KIA00793
; LOCATION: (1)..(3997)
; OTHER INFORMATION: GenBank accession No. AB018336
US-10-802-432-16

Query Match 33.1%; Score 1037.4; DB 9; Length 3997;
Best Local Similarity 60.2%; Pred. No. 6,76-302;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

Qy 21 GCCGACCCGAGATCACTGAGGAGGCGCGGAAATTCGGGGATCAGTACTTTGAAAG 80
Db 150 GCAGACTGCGAGGATGCGCTTGGGTGCCAGACCCCTGTGGAGTTAGCACCCTTGAGCC 209

Qy 81 TGGCAGAAAGCCGCCCAACACCTTCAGGAAATCTGTGCTCAATCAAAATTCAGATGT 140
Db 210 TGGGCAAGACTCTTGTGCCAGAAATGCAAGAAAGACCTGACCTCAGATTAAGCTGT 269

Qy 141 GGAATGACCCGAGGAGGATTTGAAATTCACAAAGAGCTCTGGAGGTGCTGTGA 200
Db 270 GGAACAACCCATGGAATATTTGACATTTGAGCCCTAAATGCGATGGCCAGGTATTAAGT 329

Qy 201 TGCAGTTTGCACACACTCAACCTCGTGAAGGATATTTTGGGCTCGAGTTTCTCTGA 260
Db 330 ACAAGTGTGAAAGCTTTAAACCTGTAGATGACTTCTGGGATGAGATTTCAAAA 389

Qy 261 TCACAAAAGATCAAGGTGTGCTGATCTCTAAACCCATTGTGAAACAGATTAGAA 320
Db 390 TACTCAGTCTACTGATTTGGCTTGAACCTATGAACCCATCATTTAGCAAAATAGAA 449

Qy 321 GCCAAAGACGTTGTTGTTAAATTTGTGTAATTTCTTCCGCTGACACACAACT 380
Db 450 GCCAAAGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 509

Qy 381 CCAAGAAAGATCAAGGATACCTGTTGCGCTGCGAGTGAAGAGAGAGAGAGAGAGAG 440
Db 510 ACAAGAAAGATTAACAAGATATTTGTTGCTTGCACCTTAAGAGAGAGAGAGAGAGAG 569

Qy 441 CAGGTGACGTTGATGACACAGCGAGCTCTTGTGATTTCAACATTTGCAATTTGA 500
Db 570 GCGTTTGACCTGTGCTGACACAGCGGCGCTTCTCAGCTCCCATCTCTGCGAGTCCA 629

Qy 501 GATTGGGATTTTGTATGAAAGCTTTGACAGAGAGCACTTACGAAAAATTAATACATCC 560
Db 630 AATAGAGATTAAGATGAAAGCGCTGACGAGAGCACTCAAAAGTGAACAGATATTGGCC 689

Qy 561 TCAGCAAGACGACTAGAGCAAAATCTGAAATTTCAACATTAACCAATTTGAGCAAA 620
Db 690 TGGCCAGCAGCAGCTGCTTGAAGAAAGATTAAGAAATTTCAACAGAGAGAGAGAGAG 749

Qy 621 ACCAGCAAGATCAAGATTTCCAGCTCTTGAAGATTTGCCGTGAGTGAAGATGATGAA 680
Db 750 ACCTGCTGAGTGGATTTCCAGGTGCTGAAATTTGCTGAAAGTTGAAATGTAAGGCA 809

Qy 681 CCGTTGACCCGCGCAAGAGCAGGAAAGGCAAGATCAATCTGGCGTTGCCAACAC 740
Db 810 CAGATTTCAATGAGCTTCTGACAGGAGAGAACCAAGATTTCACTGGCAGTTTCCCAT 869

Qy 741 GGAATTTCTAGTGTTCAGGGTTTCAATGATCAATGCTTCACTGAGGCGCAAGTGC 800
Db 870 GGGTGTACTGTGTTCAGAGGACACCAAAATCAACACTTTCAGTGTGCAAGGTCCG 929

Qy 801 GAACTGAGCTTCAAGAGAGAGCGCTTTCATCAAGCTCCGCGCAGATGCCAATAGTC 860
Db 930 TAAACTAGCTTCAAGAGAGAGAAATTTCTATCAAACTTCATCAGAGGTTGATGACC 989

Qy 861 GTACCAAGATACCTTGAATTCGATGAGCAGTGGGATTTTCGAAGCTTTTGGAA 920
Db 990 TTACAGAGACATTAAGAAATTTTGTGGGTAGAGATGAATGTAAGAACTTTGGAA 1049

Qy 921 AATCTGTGTAACATCATGCTTCTTGAATTTTGAAGAGCCCAACCAAGCCCA 980
Db 1050 GATTTGTGAGATATCAACCTTTTGAATTTTGAACCAACCTTAAGCCAAAGCAA 1109

Qy 981 GCCGTCCTTTTAAAGCGGGGTCATCATTTGCTGATGCTGATGCTGATGATGATGAT 1040
Db 1110 AGCGCTTCTTCAAGCGGGGCTCTCTTCAATTAAGTGAAGAACTCAGAAACAAT 1169

Qy 1041 TCTGCACTATGTTAAAGAGAGAGACATTAAGAGTGCAGTTTGAAGAGAGAGCA 1100
Db 1170 AGTAGATTTATTAAGACAGTGAATGAAGAAATTCATATGAAGAGAGAGCAAC 1229

Qy 1101 GATTCATTTATCTGAGAGCTTCTTCAAGCTTCAAGATCAAGATTTGGAATGCTGGA 1160
Db 1230 GACCCACAGTCC-----GTTGAGCTCTGACAGACCTTAACCTTAACCAA 1271

Qy 1161 GCACTCAGCAGAGACACAGCTTAATTTGAAGAAAGTCCGAAATTCAGAGGGCCA 1220
Db 1272 ACAGACATCTCATTTCCCGAGGAGATTTGAGAGACTCTGCTCCCAATCTTCAAGATGC 1331

Qy 1221 GAGCTCCGCGAGAGAAAGAACCAAGAGTTTCCGCGGAGACCGGGGTGCAACCCGAG 1280
Db 1332 CTTTACTGCTCTCTCTCTCACTGTGCTCTCTCTGAGCTTCAAGAGTTTAAGACAG 1391

Qy 1281 CCTGCGCGAGAGAGAGCCCGCGGTTAAACAAGCGAGAGAGAGAGAGAGAGAGAGAG 1340
Db 1392 CAGCAGCTCCTCAAGATCCCAAGTTTCTTACAGATGTCAGAGTCAAGAGAGAGAG 1451

Qy 1341 CACGAGAAAGAGAGAGAG--TCTTAAGATGAGACCCAGCAGATTAACCTCAAGCC 1397
Db 1452 CAGTGAAGCAGTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1511

Qy 1398 CCGGAG 1457
Db 1512 ACTCAGAGCTGTGAG 1571

Qy 1458 CTCGAG 1517
Db 1572 GAGCCCTGAGTCTGAG 1622

Qy 1518 CAG 1577
Db 1623 ACAGAGCTATACCCCACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1682

Qy 1578 CAGTGAAGATGAG 1637
Db 1683 CTCGAGAG--GAGCCAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1739

Qy 1638 GGAAGTCTTACACAG 1697
Db 1740 AGAGATTTCTGCTAAG 1799

Qy 1698 TCAGAGCAAGTGAAG 1757
Db 1800 CCGCAGCGAGAGTGAAG 1859

Qy	1758	GAATTTTGAACCTTTGACAAATTTCTATCTAAATTTTCTCAAGAAATTTGACAAACACT	1817
Db	1860	CAACATTCATCCCATCTTAAGATTCCACAGAGCTTCTCGCGCGAGGTGAGCAGAGCTT	1919
Qy	1818	TGCGCCTGTGGAAAGGCGCTCAAAATGCCCAATTCAGAG---ATTACAAAGATTCGGCGA	1874
Db	1920	GGCATTCTGGGAAGGGCCCTCCAAAGCCACCAAAAGGCACTCATCAAGATTCGGGA	1979
Qy	1875	TGTCA TGTGAAGAACTTCAAGGGATGAAGCAGCTGGCGGCTCACTGTGAAGACAG	1934
Db	1980	CATCTCGCTCAGAAACATGCGCAGATTAAAGAGTTTACAGACTCTTCCAAAGACATGA	2039
Qy	1935	CGAGGCTTTGAGAGCCCTTGAGAAATGGAATCAAGAGCTCCGGCGGCTGGAACCTCTG	1994
Db	2040	CGAGGCTCTTACAGAACTTGAAAAGGCTTACCAAGCTGTAAAGATTGAGGCAAGTGT	2099
Qy	1995	CAGAAGCTTTGAGCTGACAGAAGTGTGTTACCTACCGCTCAACCTTCTCTGCGGCG	2054
Db	2100	CAGAAGTTTGAAGCTGACAAAGGTCTGTCTACTTGCCTTCAACAGTTTCTGTGAAGCC	2159
Qy	2055	AATGCA CCGGCTCATGCACTAACAGAGGTCTTGAGCGGCTGTGCAAAACACACCCGCG	2114
Db	2160	CATCAGAGGGCTGTGCTGCACTTACCGGCTGTGCTGGCGCGGCTATCGGACATTCAGCC	2219
Qy	2115	GAGCCACCGGCATCTTCAGGGAATGCCAGGCCGCTTTGGCAGAGATCACAGAGATGTGTGC	2174
Db	2220	CGGGCACCAATGACTACGCTGACCTGCATGCCCTGTAAAGCATCACAGAGGTGACCA	2279
Qy	2175	ACAGCTCCA CGGTACGATCAATGAGTGAATTTCCAGAAAGCTGACGAACTCAAGAA	2234
Db	2280	CACACTACAGCAATTTCTATCCGGCTGGAAGACTGCAAGAGCTTACGAGACTTCACCG	2339
Qy	2235	AGATTTGATTGGCATTTGACAACTTTGTGGTTCGGGAAAGGAGTTCACTCGTCTGGCAG	2294
Db	2340	GGACCTGGTGGGCAATAGAGAACTCATTTGTCTCTGGCAGGAGTTCACTCGTGAAGGCTG	2399
Qy	2295	CCTCAGCAAGCTCTCGGGGAAAGGGCTCCAGACGCAATGTTCTTCTGTTCACAGACT	2354
Db	2400	CTTTCACAAGCTCACCAAGAAAGGCGCTGCAGCAGAGAGATGTTTCTGTTCTCAGATAT	2459
Qy	2355	CGTGCATTCAGAGCCGGGGGCTGACGGCTCCAAATCAGTTTAAAGTCACGGGCACT	2414
Db	2460	GTTTGTCTGTCACAAAGAAAGGTTGACAGGACACAGCCACTTCCGGATCCGGGGCTCTCT	2519
Qy	2415	CCCGCTCTATGGCATGACGATTGAGAGAGCAGAACAGAGTGGGGGTGCCCTCCTGCT	2474
Db	2520	TCCCTCCAAAGGCAATGCTGTGTGAAGAAAGTATACAGATGTGTGTTCCACACTGTTT	2579
Qy	2475	GAACCTCCGGGGCCAGCGGCACTCATCTCGTGGCCGCGCACTTCTCGGTCCAGATGGA	2534
Db	2580	CACCAATCTTCGGCGGCTCAGAAAACAATCGTGTGGCACACAGCACTCGGTGAGAAAGA	2639
Qy	2535	GAAATGGGTGTGAGGACATCCCAATGGCAATGACCTGGCGGAGAAAGACACACAGCCCGC	2594
Db	2640	GAAATGGAGTGTGAGACTGTAACTCCGCAATCCAGACAGCAAGATGGCGGTGACACGGC	2699
Qy	2595	CCCTGAGTTCTGTGACAGCA-----GCCGCCCTGACAAACAAGTCCCTGTATGAAGCAC	2648
Db	2700	CCCTGTGACTGCCAGGCGGCACTGTGTGCACTGTCCCCCAATGCCCAACAGATATC	2759
Qy	2649	CGCGGCTGACAGAGTCAAGAGATGACTGAAGCCCTTCGGGCAATCGCTGAGAGCGCA	2708
Db	2760	TCGTG---GAGCAGGAGTCAAGAGATGATCTCGGGGTGTCCGCAAGCTCCCTGAGGGGCA	2816
Qy	2709	GGCCCCGCA CCGCGGCAACACAATGTGTGACTGTGTGCTGGCACCGGCAACACAGCTCTC	2768
Db	2817	TGGCCAGCACCGGGGCCAACACAATGACGTGTGTGTGTATCCGAAACACAGGTGTCT	2876
Qy	2769	CATGTGTGACTTCAGCATCGCAGTGGAGAAATCAGTTGTCTGGAAACTGTGAGGAAAT	2828
Db	2877	CAGGGCAGACACAGTGCAGCTGTGTGAGAACCAAGCTTTCAGGATATCTGTCTTAAGAAATT	2936
Qy	2829	CAAAAACAGCAACGGGTGGCAGAACCTGTGGTGTGTTCAAACTTCTGCTGTTCTT	2888

Accession	Sequence	Position
Db	2937 CAAAACAGTCATGGCTGGCAGAGACTCGGGTCGCTTTACCACTTCTGTTTGTCTT	2596
Qy	2889 CTACAAATCACCCAGGACAATCATCCCTTSCCAGCCTGGCCTCTGCTGGCTACTCGCT	2948
Db	2997 CTACAAAATCTATCTAGATGACTACCCACTGGCCAGCCCTCCGCTCTGGGCTACAGCGT	3056
Qy	2949 CACCATCCCTCTGATGCCGAGAACATCCAGAAAGCTACGTTCMAAGTCGACTTCAA	3008
Db	3057 GAGCATCCCCCAGGGGAGCCGATGGCATACAAAGACTATGTTTTCAAGCTCAGTTCAA	3116
Qy	3009 GTTCCACGCTACTACTTTCAGGGGGGAAAGCAGTACAAGTCGAAAGGTGATGAGGT	3068
Db	3117 ATCCACAGCTACTCTTCTCCGGGCTGAGAGCAAGTACACATTTGAAAGGTGATGAGGT	3176
Qy	3069 GATCCGACGTGCCACAGCTCTGCTCCGGAATCC	3103
Db	3177 GATCCAGGGGGCCAGCAGCTACGCCGGGAGGGGCC	3211

```

RESULT 8
US-09-764-868-51
; Sequence 51, Application US/09764868
; Patent No. US2002016871A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 1718
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-51

```

Query Match	12.8%	Score 402.4	DB 3	Length 1718
Beet Local Similarity	67.1%	Pred. No. 6.2e-110		
Matches	603	Conservative	0	Mismatches 286; Indels 9; Gaps 2
Qy	2212	CAGAGCTGCACGAACTCAAGAAATTTGATTTGGCATTCACATTTGTGTGCTCCGGA	2271	
Db	13	CAGAAGCTACAGAGCTGCAGCGGGAACCTGGTGGGCATAGAAACCTCATGTGCTCGGC	72	
Qy	2272	AGGAGTTCAATCCGTCGTCGGGACCTCAGCAAAGCTCTCGGGAAAGGGCTCCAGACGCG	2331	
Db	73	AGGGAGTTCAATCCGTGAGGGGCTGCTTCACAAAGCTCACAAAGAGGGCTGACAGAGAG	132	
Qy	2332	ATGTTCTTCCTGTTGAAAGAGCGTCCCTGCAATTCAGAGCGGGGGGCTGACGGCTCCAT	2391	
Db	133	ATGTTTTTTCGTGTTTCAGATATGTTGCTGTAACAAGAAAGAGTGTGAGGAGCACGAC	192	
Qy	2392	CAGTTTAAAGTTCACGCGACGCTCCCGCTCTATGCGATGACGATTGAGGAGAGCGAAGAC	2451	
Db	193	CACTTCCGGATCCGGGGGCTCCTTCCCTCCCAAGGATGCTGTGGAAGAAATGTGATTAAC	252	
Qy	2452	GAGTGGGAGGTGCCCCACTGCTGACCCCTCGGGGCGACGGCGCATCATCTGTGGCC	2511	
Db	253	GAGTGGTCTGTTCACACATGTTTCCATCTACGGCGGCTCAGAAAAAATCTCGTGTGGCA	312	
Qy	2512	GCGAATTCCTCGTCCGAGATGAGAGAGGGGTTGAGGACATCAGATGGCCATTGACCTG	2571	
Db	313	GCGACACATCTCGCTGGAGAAAGAAAGAGATGCTGGAACCTGACCTCCGGATCCAAAGCA	372	
Qy	2572	GCGAGAAAGACGACGAGCCGCCCTGTGATTCCTGTGCAGCA-----GCCGCCCTGAC	2625	
Db	373	GCGAAGATGGGGGTGACACGGGCCCTGTGACCTGACGCGGACCTGTGTGCACTCGTCCC	432	
Qy	2626	AACAAATCCCTGATGAAGCACCGCGGCTGACCAAGATCAGAGATGACCTGAGCGCC	2685	


```
/ Publication No. US20050064454A1
/ GENERAL INFORMATION:
/ APPLICANT: Avalon Pharmaceuticals, Inc.
/ TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
/ FILE OF INVENTION: Signature Gene Setc
/ FILE REFERENCE: 689290-189
/ CURRENT APPLICATION NUMBER: US/10/843,641A
/ CURRENT FILING DATE: 2004-05-12
/ PRIOR APPLICATION NUMBER: US/09/873,367
/ PRIOR FILING DATE: 2001-06-05
/ PRIOR APPLICATION NUMBER: US/09/954,531
/ PRIOR FILING DATE: 2001-09-18
/ PRIOR APPLICATION NUMBER: US/09/954,456
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US/09/962,436
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US/09/962,832
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US/09/964,824
/ PRIOR FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: US/09/967,768
/ PRIOR FILING DATE: 2001-09-28
/ PRIOR APPLICATION NUMBER: US/09/968,007
/ PRIOR FILING DATE: 2001-10-02
/ PRIOR APPLICATION NUMBER: US/09/969,347
/ PRIOR FILING DATE: 2001-10-02
/ PRIOR APPLICATION NUMBER: US/09/969,708
/ PRIOR FILING DATE: 2001-10-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 8447
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 6249
/ LENGTH: 506
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-843-641A-6249

Query Match          11.4%; Score 358.4; DB 10; Length 506;
Best Local Similarity 99.7%; Pred. No. 6,7e-97;
Matches 359; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2776 GACTTCAGCATCGCAGTGGAGAAATCAAGTGTCTGAAAACCTGCTGAGAAATTCAAAAC 2835
DB 506 GACTTCAGCATCGCAGTGGAGAAATCAAGTGTCTGAAAACCTGCTGAGAAATTCAAAAC 447
QY 2836 AGCAAGGGTGGAGAAAGCTGTGGTGTGTTCACAACTTCGCTGTTCTTCAAAA 2895
DB 446 AGCAAGGGTGGAGAAAGCTGTGGTGTGTTCACAACTTCGCTGTTCTTCAAAA 387
QY 2896 TCACACACAGGACATATCCCTTGGCAGCTGCTCTGCTGAGCTACCTCAACATC 2955
DB 386 TCACACACAGGACATATCCCTTGGCAGCTGCTCTGCTGAGCTACCTCAACATC 327
QY 2956 CCCTCTGAGTCCGAGAAATCCAGAAAGACTACGTGTTCAGCTGCACTTCAAGTCCAC 3015
DB 326 CCCTCTGAGTCCGAGAAATCCAGAAAGACTACGTGTTCAGCTGCACTTCAAGTCCAC 267
QY 3016 GCTTACTACTTCAGGGCGGAAAGCGAGTACAGCTTGAAAGGTGATGGAAGTATCCCG 3075
DB 266 GCTTACTACTTCAGGGCGGAAAGCGAGTACAGCTTGAAAGGTGATGGAAGTATCCCG 207
QY 3076 AGTGCCACAGCTCTGCTCGCGACCCCAAGCTGTTAGCCACAAAGATCTTTGTGTAT 3135
DB 206 AGTGCCACAGCTCTGCTCGCGACCCCAAGCTGTTAGCTACAAAGATCTTTGTGTAT 147

RESULT 12
US-09-960-352-154
/ Sequence 154, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Nengbing
```

```
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathialagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
/ FILE REFERENCE: 16511.006/37-21 (10298) C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112
/ SEQ ID NO 154
/ LENGTH: 379
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 01-LIB34-020-Q1-E1-A9
US-09-960-352-154

Query Match          9.5%; Score 297; DB 3; Length 379;
Best Local Similarity 89.7%; Pred. No. 2.2e-78;
Matches 331; Conservative 0; Mismatches 35; Indels 3; Gaps 1;

QY 1587 TGAGGGCCGAGAGAGATTCCTCAACTGATTAAGGCTACTTCATAGCTAAGAGTGTG 1646
DB 11 TGGGGCCCGAGAGAGATTCCTCAACTGATTAAGGCTACTTCATAGCTAAGAGTGTG 70
QY 1647 TACCAACGAGCGAATATCTGAAGATCTGAAGTATCACTTGTGCTTTCAGAGCAC 1706
DB 71 CACCACTGAGAGACATATCTGAAGATCTGAAGTATCACTTGTGCTTTCAGAGCAC 130
QY 1707 AGTGAGCAAGAGAGAGAGATTCCTCAAGAGCTGAAAGATCTCATATTCCTCCAAATTTGA 1766
DB 131 AGTGAGCAAGAGAGAGATTCCTCAAGAGCTGAAAGATCTCATATTCCTCCAAATTTGA 190
QY 1767 ACCCTTGCAAAATTTCTAATATTTCTCAAGAAATTTGAGCAAGCACTTGCCTGTG 1826
DB 191 ACCCTTGCAAAATTTCTAATATTTCTCAAGAAATTTGAGCAAGCACTTGCCTGTG 250
QY 1827 GGAAGCCGCTCAATATGCCAATCA---GAGATTACCAAGATGCGGATGTCTGCT 1883
DB 251 GGAAGCCGCTCAATATGCCAATCA---GAGATTACCAAGATGCGGATGTCTGCT 310
QY 1884 GAAAGAACTTCAGGGGATGAACACCTGGCGGCTACCGCTGGAAGACAGAGGCGCTT 1943
DB 311 GAAAGAACTTCAGGGGATGAACACCTGGCGGCTACCGCTGGAAGACAGAGGCGCT 370
QY 1944 GGAAGCGCT 1952
DB 371 GGAAGCGCT 379

RESULT 13
US-09-960-352-156
/ Sequence 156, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Nengbing
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathialagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
/ FILE REFERENCE: 16511.006/37-21 (10298) C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112
/ SEQ ID NO 156
/ LENGTH: 400
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 01-LIB34-020-Q1-E2-A9
US-09-960-352-156

Query Match          9.3%; Score 292; DB 3; Length 400;
Best Local Similarity 87.5%; Pred. No. 7.3e-77;
Matches 343; Conservative 0; Mismatches 45; Indels 4; Gaps 2;
```


This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2006, 09:18:29 ; Search time 32.413 Seconds

(without alignments)
6824.209 Million cell updates/sec

Title: US-09-555-342b-1_COPY_49_3183

Perfect score: 3135
Sequence: 1 atgggggaatcagacagag.....acaagagctctgtgtat 3135Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 224314 seqs, 35277956 residues

Total number of hits satisfying chosen parameters: 448628

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA New:*

- 1: /EMC_Celerra_SIDS3/prodata/1/pubpna/US09_NEW_PUB.seq:*
- 2: /EMC_Celerra_SIDS3/prodata/1/pubpna/US06_NEW_PUB.seq:*
- 3: /EMC_Celerra_SIDS3/prodata/1/pubpna/US07_NEW_PUB.seq:*
- 4: /EMC_Celerra_SIDS3/prodata/1/pubpna/US08_NEW_PUB.seq:*
- 5: /EMC_Celerra_SIDS3/prodata/1/pubpna/PCT_NEW_PUB.seq:*
- 6: /EMC_Celerra_SIDS3/prodata/1/pubpna/US10_NEW_PUB.seq:*
- 7: /EMC_Celerra_SIDS3/prodata/1/pubpna/US11_NEW_PUB.seq:*
- 8: /EMC_Celerra_SIDS3/prodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	192.4	6.1	2867	6 US-10-511-937-344	Sequence 344, App
2	56.8	1.8	3953	6 US-10-473-173-56	Sequence 56, App
3	41.8	1.3	6466	6 US-10-488-619-2736	Sequence 2736, App
4	41.4	1.3	37426	6 US-10-473-173-32	Sequence 32, App
5	40.4	1.3	3708	7 US-11-325-764-1	Sequence 1, App
6	38.4	1.2	2367	6 US-10-505-928-744	Sequence 744, App
7	36.4	1.2	2711	7 US-11-145-307A-189	Sequence 189, App
8	35.6	1.1	1333	6 US-10-505-928-380	Sequence 380, App
9	35.4	1.1	4773	6 US-10-511-455-1	Sequence 1, App
10	34.4	1.1	909	6 US-10-473-173-50	Sequence 50, App
11	34.2	1.1	4450	6 US-10-505-928-275	Sequence 275, App
12	34	1.1	4234	7 US-11-312-958-25	Sequence 25, App
13	34	1.1	4573	6 US-10-505-928-409	Sequence 409, App
14	34	1.1	4585	6 US-10-505-928-47	Sequence 47, App
15	33.8	1.1	555	7 US-11-217-529-174434	Sequence 174434, App
16	33.8	1.1	2121	6 US-10-196-749-355	Sequence 355, App
17	33.8	1.1	2121	7 US-11-101-316-117	Sequence 117, App
18	33.8	1.1	4740	6 US-10-511-937-380	Sequence 380, App
19	33.8	1.1	8833	6 US-10-505-928-785	Sequence 785, App
20	33.6	1.1	885	7 US-11-217-529-173230	Sequence 173230, App
21	33.4	1.1	5411	6 US-10-504-120-5	Sequence 5, App
22	33.4	1.1	5543	6 US-10-504-120-6	Sequence 6, App
23	33	1.1	2294	6 US-10-511-937-513	Sequence 513, App
24	32.6	1.0	819	7 US-11-024-544A-119	Sequence 119, App
25	32.6	1.0	819	7 US-11-024-545-47	Sequence 47, App

26	32.6	1.0	819	7 US-11-185-301-35	Sequence 35, App
27	32.6	1.0	819	7 US-11-190-750-102	Sequence 102, App
28	32.6	1.0	819	7 US-11-251-466-21	Sequence 21, App
29	32.6	1.0	819	7 US-11-254-173-35	Sequence 35, App
30	32.6	1.0	819	7 US-11-264-784-22	Sequence 22, App
31	32.6	1.0	12649	7 US-11-024-544A-110	Sequence 110, App
32	32.6	1.0	12649	7 US-11-024-545-38	Sequence 38, App
33	32.6	1.0	12649	7 US-11-185-301-27	Sequence 27, App
34	32.6	1.0	12649	7 US-11-190-750-93	Sequence 93, App
35	32.6	1.0	12649	7 US-11-251-466-12	Sequence 12, App
36	32.6	1.0	12649	7 US-11-254-173-26	Sequence 26, App
37	32.6	1.0	12649	7 US-11-264-784-114	Sequence 114, App
38	32.4	1.0	2811	6 US-10-511-937-648	Sequence 648, App
39	32.4	1.0	4050	6 US-10-511-937-663	Sequence 663, App
40	32.4	1.0	4593	6 US-10-524-021-8	Sequence 8, App
41	32.4	1.0	5994	6 US-10-505-928-701	Sequence 701, App
42	32.2	1.0	550	6 US-10-488-619-2469	Sequence 2469, App
43	32.2	1.0	627	6 US-10-488-619-2468	Sequence 2468, App
44	32.2	1.0	1086	7 US-11-217-529-1630	Sequence 1630, App
45	32.2	1.0	2040	7 US-11-217-529-5680	Sequence 5680, App

ALIGNMENTS

RESULT 1
US-10-511-937-344
Sequence 344, Application US/10511937
GENERAL INFORMATION:
APPLICANT: EXPRESSION DIAGNOSTICS, INC.
APPLICANT: Wohlgemuth, Jay
APPLICANT: Fry, Kirk
APPLICANT: Woodward, Robert
APPLICANT: Ly, Ngoc
APPLICANT: Prentice, James
APPLICANT: Morris, Macdonald
APPLICANT: Rosenberg, Steven
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
FILE REFERENCE: 506612000104
CURRENT FILING DATE: US/10/511, 937
CURRENT FILING DATE: 2004-10-19
PRIOR APPLICATION NUMBER: PCT/US2003/012946
PRIOR FILING DATE: 2003-04-24
PRIOR APPLICATION NUMBER: US 10/131, 831
PRIOR FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: US 10/325, 899
PRIOR FILING DATE: 2002-12-20
NUMBER OF SEQ ID NOS: 3117
SOFTWARE: PatentIn version 3.2
SEQ ID NO 344
LENGTH: 2867
TYPE: DNA
ORGANISM: Homo sapiens
US-10-511-937-344

Query Match 6.1%, Score 192.4, DB 6, Length 2867,
Best Local Similarity 52.8%, Pred. No. 7.5e-43,
Matches 490, Conservative 0, Mismatches 426, Indels 12, Gaps 3;

QY	117	CGGCTGATCAAAATGCTGATGACCCAGGAGCATTTGAGTCCACAAG	176
DB	798	CATGACCTGAGAGCTTCTTGTGATGACAGTTTGAATGTGTGGAGACAC	857
QY	177	AGCTCTGGAGAGTCTGCTGATGACAGTTTGCAACCTCAACTCGTGAAGTGA	236
DB	858	TGCTAAGGACAGATTTCTTAAAGAGATGTAGATCTCAATCTTTGGAGACAA	917
QY	237	CTATTTTGGCTTGAATTCCTGATCAAAAAGATCAGCGTGTGCTGATCTCTAA	296
DB	918	CTATTTTGTGATGACATTTGGATTAACCACTTGAAGATGCTGATTCGCCAA	977

QY	297	1	CCCATTTGTGAAACGATTTAGAGGCCAAGACGCTGTGTGTAAGTTGTGTGTAATTT	356
Db	978	AGAAATAAAAAAGCGGTTTCGGG--TGTCCTTGGAAATTTTACATTTAATGTAAGTT	1034	
QY	357	CTTTCGCGCTGACCAACACAACTCCAGAAGAAAGTCAACAAAGTACTCTGTTCGCGCTGCA	416	
Db	1035	TTATTCACCTGACCCAGACAGCTTAAAGAAACATTAACAAATATTATTATGTCTTCA	1094	
QY	417	GGTGAAGCAGGACTTGGCTCAAGGCAAGTTGACGTGTATGACACACGCGCAGCTCTCTT	476	
Db	1095	GCTTCGGACGACATAGTTGGCAGGACGTCGCTGCTTCCTTTGGCAACTTTCAGATTAAT	1154	
QY	477	GATTTCAACAATGTGCAATCTGAATTTGGGGATTTTGTATGAAGCTTGGACAGAGACA	536	
Db	1155	AGGTTCTTAAACCATTCAGTCTGAATCTGGAGACTACGCCAGAACTTCACAGCGGTGGA	1214	
QY	537	CTTAGCAAAAAATATAATACATAC-----TCAGCAGACGCACTAGAGGACAAATCGT	590	
Db	1215	TTATGTATGTATTTTAACTGGCCCCCAATACAGACAAAGAACTTGAAGAAAGTCAAT	1274	
QY	591	GGAATTTGACCAATTAACAACATTGGACAAACACACAGCAGAAATCAGATTTTCAGCTCCTAGA	650	
Db	1275	GGAATGTCATAGTATCATACAGGTCATGATCTCAGCTCAGGCTGACTTGAATTTCTTGA	1334	
QY	651	GATTGCCCCGTGGGCTAGAGATGTATGGAATCCGGTTGACCCGGGCCAAGAACAGGAAG	710	
Db	1335	GAATGCCAAAAAGTTGTCATGTGTATGAGAGTTGATCTTCAATAAGCAAAAGACCTTGGAAAG	1394	
QY	711	CACGAAATCATCTGGCCGCTGGCCAAACACGGGAAATTCAGTAGTTTCAGGCTTTCACATTA	770	
Db	1395	AGTAGATATCATCTTAGGTGTCTGTCTTAGTGCCCTTCTGGTTTACAAAGATTAAGCTGAG	1454	
QY	771	GATCAATGCTTTCAACTGGGCCAAGAGTGGGAAAGCTGACCTTCAAGAGGAAACGCTTCT	830	
Db	1455	AATTAAACGGCTTCCCTGGCCCCAAAGTGTGAAAGATTTCTTAAACGTAAGACTTTT	1514	
QY	831	CATCAAGCTCGGGCCAGATGCCAATATAGTCGTATCCAGATACTCTTGAATTCCTGATGCG	890	
Db	1515	CATCAAGATTCGGCCCTGGAGAACAAAGACGATATGAAAGTACATTCGGAATTTAAACTTCC	1574	
QY	891	CAGTCGGGATTTCTGCAAGTCCCTCTGGGAAATCTGTGTGAACATCATGACCTCTTATG	950	
Db	1575	CAGTTTACCGAGACGCTAAGAAATTTATGAAAGTCTGTGTGAAACATCACAGCTTTTTCAG	1634	
QY	951	ACTTTTGAAGACCCAAACCAAGCCCAAGCCGCTCTCTTTAGCCGGGGGTCAATCAT	1010	
Db	1635	A---TTGACATCTACAGACACCATTTCCAAAGCAAAATTTCTTGGGCTAGATCCAAAT	1691	
QY	1011	TCGGTTCAAGTGTGGGACTCAGAAAGCG	1038	
Db	1692	TCGATACAGTGGCCGACTCAAGCTCAG	1719	

```

RESULT 2
US-10-473-173-56
? Sequence 56, Application US/10473173
? Publication No. US2006008823A1
? GENERAL INFORMATION:
? APPLICANT: VAN ANDEL INSTITUTE
? TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
? TITLE OF INVENTION: Carcinoma: Prognosis and Drug Target Identification
? FILE REFERENCE: 38345-170094
? CURRENT APPLICATION NUMBER: US/10/473,173
? CURRENT FILING DATE: 2003-09-29
? PRIOR APPLICATION NUMBER: US 60/279,411
? PRIOR FILING DATE: 2001-03-29
? NUMBER OF SEQ ID NOS: 498
? SOFTWARE: Patencin version 3.2
? SEQ ID NO 56
? LENGTH: 3953
? TYPE: DNA
? ORGANISM: Homo sapiens
US-10-473-173-56

```

Query Match	1.8%;	Score 56.8;	DB 6;	Length 3953;
Best Local Similarity	46.6%;	Pred. No. 6.6e-06;		
Matches 218; Conservative	0;	Mismatches 247;	Indels 3;	Gaps 1;

QY	130	ATCCGATGCTGGATGAGACACCAGGAGATTGAAGTTCACCAAGAGAGCTCTGGGAA	169
Db	284	ATCCACTGCTTAATATACGAGTGTTTGGAGTTACACCTCTCCGTGAGAGACACTGGCCAG	343
QY	190	GTGCTGCTGGATGTCAGTTTGGACCAACCTCAACTCTGTGAAAGTGACTAATTTGGCCTC	249
Db	344	GAAGAGCTCGAGGCGGTGGCCAGAGAGCTGAGACTGCGGGAAGTCACTTACTTCAACCTC	403
QY	250	GAGTTTCTGATCACAAAAAGATCAAGGTGGAGTGGATCTCTCTAAACCATTTGGAA	309
Db	404	TGGTACTACAAACAAACAAATACGCGCGGTGGTATTTGGAAAACCTTTGAAGAA	463
QY	310	CAGAT---TAGAAGGCCAAAGCACGCTGTGTTTAAATTGTGGTGAATTCCTTTCCGCT	366
Db	464	CAGCTGGATTAATAATATGCAATTTGGAACCTACCGTCTAATTTTGGAGTGTGTTTATGTGCT	523
QY	367	GACCAACAACAATCTCCAAAGAGAATCTCACAGATACCTTTCGCGCTGCAGGTGAAGCAG	426
Db	524	TCAGTTTCTCAGCTCCAGAGAGAGATTACAGATATCATATATATTCGAACAACGAAAGAA	583
QY	427	GACTTGGCTCAAGGAGGTGACGTATATGACACAGGCGACGCTCTTGATTTTCAAC	486
Db	584	GATATCTTGGAGGAATATTCCTTTGTAACCTTAAGAACAAAGCAATTAAGCTAGCAGGCTTA	643
QY	487	ATTGTGCAATCTGAGATTGGGATTTTGAATGAAGCCTTGGACAGAGAGCACTTAAGCAAA	546
Db	644	GCTGTTCAGGCGAATTTTGTGTACTTTGATCAGTATGAATCCAGGACTTTCTTCAGAAA	703
QY	547	AATAAATACATCTTCAGCAAGACCACTAGAGGACAAATGCGTGA	594
Db	704	TTTGCTGTGTTCTCTGTGGAGTGTTCACAGATGAAGAAAAAGTATTGGAA	751

```

RESULT 3
US-10-488-619-2736
; Sequence 2736, Application US/10488619
; Publication No. US2006009578A1
; GENERAL INFORMATION:
; APPLICANT: Greentree, Winner and Sullivan, P.C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations &
; TITLE OF INVENTION: Physiological Conditions, And Genotyping Arrays
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US-10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: Patentin version 3.1
; SEQ ID NO: 2736
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-2736

```

Query Match	1.3%	Score 41.8	DB 6	Length 646
Best Local Similarity	49.8%	Pred. No. 0.028		
Matches	106	Conservative	0	Indels 0
				Gaps 0
QY	2007	GCTGACAGAGGTGTTGTTACCTPACCGCTCAACACCTTCTCTCTGGGGCCACTGCACCGGCT	2066	
Db	73	GCTGCAAGAGATGATGACATCTGGCTGGATGGCTTTCTGCTGACCCCGGTGCAGAGAT	132	
QY	2067	CATGACCTACAGAGAGTCTCTGAGCGCGCTGTGCAACACCCACCGCGGACCGCCA	2128	
Db	133	CTGCAAGATACCCCTCCAGCTGGGGGAGCTGCTCAAGTACACACACCCCGACAGGGA	192	
QY	2127	CTTACAGGACCTGCCGAGCCGCTTTGGCAGAGATACAGGAGATGTGGCACAAGCTTCACGG	2188	
Db	193	CTTTAAGATATTGAAGCTGCTCTTGACAGCCCATAGAGATGTGGCCAGCTCATATACGA	252	

Qy 2187 TACGATGATCAAGATGAGATTTCAGAACT 2219
Db 253 ACGGAAACGAAGCTTGAAAACATCGACAGAT 285

RESULT 4

US-10-473-173-32
; Sequence 32, Application US/10473173
; Publication No. US2006008823A1
; GENERAL INFORMATION:
; APPLICANT: VAN ANDEL INSTITUTE
; TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
; FILE REFERENCE: 38345-170094
; CURRENT APPLICATION NUMBER: US/10/473,173
; PRIOR FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US 60/279,411
; NUMBER OF SEQ ID NOS: 498
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 32
; LENGTH: 37426
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-173-32

Query Match 1.3%; Score 41.4; DB 6; Length 37426;
Best Local Similarity 54.2%; Pred. No. 0.37;
Matches 84; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 205 GTTTCACACCACTCAACCTCGTGAAGTGATGATTTTGCGCTCGAGTTTCTGATCAC 264
Db 11233 GTGTGACAGGCGACTGGGAATCATAGAAAGTTGACTATTTTGACTGCACTTTACGGGTAGC 11292
Qy 265 AAAAGATGATCAGGTGTGGCTGATCTCTTAAACCCATTGTGAACAGATTAGAAGGCCA 324
Db 11293 AAAGTGAAGTTTATGAGCTTAAACCTGAAACCGGATCTCCACGACGATGATGGGCTA 11352
Qy 325 AAGCAGCTTGTGTTAAGTTTGTGGTGAATTCTT 359
Db 11353 GCCCTTACAGGCTTAAACTTAGATCAAGTCTT 11387

RESULT 5

US-11-325-764-1
; Sequence 1, Application US/11325764
; Publication No. US20060099640A1
; GENERAL INFORMATION:
; APPLICANT: Jegla, Timothy James
; APPLICANT: Witzel, Julie Dickson
; APPLICANT: IChagen, Inc.
; TITLE OF INVENTION: S102 and S104, Novel Potassium Channel Proteins from
; FILE REFERENCE: 018512-006810US
; CURRENT APPLICATION NUMBER: US/11/325,764
; CURRENT FILING DATE: 2006-01-04
; PRIOR APPLICATION NUMBER: US/09/921,159
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: US 60/249,112
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3708)
; OTHER INFORMATION: human S102 potassium channel alpha subunit
US-11-325-764-1

Query Match 1.3%; Score 40.4; DB 7; Length 3708;

Best Local Similarity 58.2%; Pred. No. 0.18;
Matches 71; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 2046 CTTGCGGCACCTGACCGGCTCATGACAAAGAGTCTCGAGCGGTGTGCAACA 2105
Db 1917 CCGCTGCGCGGTGACAGATCATCGCTTCATGATGTGGCCATGACCTCGAGGGACAGA 1976
Qy 2106 CCAACCGCGGAGGACCGGACCTTCAGGAGCTGCGGAGCCGCTTTGGCAGATTCACGA 2165
Db 1977 GCACCGGCTTACGAGAGCGGCGGTGGGGGAGCAGCAAGCTGCACTGCCACGGA 2036
Qy 2166 GA 2167
Db 2037 GA 2038

RESULT 6

US-10-505-928-744
; Sequence 744, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patentin 3.2
; SEQ ID NO 744
; LENGTH: 2367
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-744

Query Match 1.2%; Score 38.4; DB 6; Length 2367;
Best Local Similarity 46.3%; Pred. No. 0.5; 146; Indels 0; Gaps 0;
Matches 126; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

Qy 1867 ATGCGGATGTCATGCTGAAGAACATTCAGGGGATGAGCACCTGGCGCTCACTGTGG 1926
Db 448 ATCTCGAGCTCGGCAAGATGATTCAGAGGTCTCCAGCACCTCAACCCCTGCTCG 507
Qy 1927 AAGCAGCGAGGCTTGAAGGCTTGAAGATGAGATCAAGAGCTCCGCGCGCTGGAG 1986
Db 508 GGAACAAAGATCTCAAGATCCATGAGAAAGGCTTCAGCCCACTGCGGAAAGTGCAG 567
Qy 1987 AACTTCTGAGAGACTTGAAGCTGCGAAGGTGTGTACTACCGCTCAACACCTTC 2046
Db 568 AAGCTTACATCTCCAAAGAACCACTGTGGAGATCCGCCCAACCTACAGCTCC 627
Qy 2047 CTGCGGCACCTGACCGGCTCATGACATCAAGAGAGTCTTGGAGCGGCTGCAAAAC 2106
Db 628 GTGAGACTCGGATCAGACAGCAACCCATCCGAGAGTCCCAAGGAGTGTTCACCGGG 687
Qy 2107 CACCGCGGAGCGGACGCTTCAGGAGCTG 2138
Db 688 CTCGGGACATGAATGATGAGATGGGCGG 719

RESULT 7

US-11-145-307A-189/c
; Sequence 189, Application US/11145307A
; Publication No. US20060094035A1
; GENERAL INFORMATION:
; APPLICANT: Arcuturus Bioscience, Inc.
; APPLICANT: Erlander, Mark G.
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Identification of Tumors
; FILE REFERENCE: 022041-002020US
; CURRENT APPLICATION NUMBER: US/11/145,307A
; CURRENT FILING DATE: 2005-06-03


```

; TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
; TITLE OF INVENTION: Carcinoma: Prognosis and Drug Target Identification
; FILE REFERENCE: 38345-170094
; CURRENT APPLICATION NUMBER: US/10/473,173
; PRIOR FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US 60/279,411
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 498
; SOFTWARE: Patent version 3.2
; SEQ ID NO 50
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-173-50

Query Match      1.1%; Score 34.4; DB 6; Length 909;
Best Local Similarity 53.8%; Pred. No. 3.5;
Matches 71; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY      1252 TCCGCGGGGAGCGGGGTCGACCCGAGCCCTGCGCGGAGGAGAAAGCCCGGGGTAAC 1311
DB      670 TAGCGGCGGCGAAGCCGAGCGCCCGGACCGGAGAAAGCGCAAGCGGGGGCGAGC 729
QY      1312 AAGCAGCGGAGCGAGCGCGCTCGCGCCCGCAGGAGGAGGAGAGTCTTAAGAT 1371
DB      730 CGCAGGCGGGGCGCGCGCGGACCCCTCTGCGGACAGAGTGAGCGCCCTTGGGGCG 789
QY      1372 AGGACCCAGCAG 1383
DB      790 CCGCACCCTCCAG 801

RESULT 11
US-10-505-928-275
; Sequence 275, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patent 3.2
; SEQ ID NO 275
; LENGTH: 4450
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-275

Query Match      1.1%; Score 34.2; DB 6; Length 4450;
Best Local Similarity 49.7%; Pred. No. 9.9;
Matches 87; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY      2058 GCACCGGCTCATCATCAAGCAGAGTCTCGAGCGGCTTGCAACACACCGCGGAG 2117
DB      2010 GCACCTGCGCAAGAAAGTACTGTCTGTCAGGCGCTTGAAAGCCCTCGGCTCAGCAGAA 2069
QY      2118 CCAGCGCGACTTCAAGGACTGCGGAGCGCTTTGGCAGAGATCAACGAGATGGTGACA 2177
DB      2070 CTTGACCTGACCTCTGTGTAACGTGAGCGACTGCTGGAATGACATGCTTGGTGGCGG 2129
QY      2178 GCTTCACGGTACATGATCAAGATGAGAAATTTCCAGAACTCAGCACTCAAG 2232
DB      2130 AGCGCAGCGCCGAGATCGTGTGTCAAGACGAGAGGCTCTGTGAGAGAAAG 2184

RESULT 12
US-11-312-958-25
; Sequence 25, Application US/11312958
; Publication No. US20060100152A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Rosenfeld, Julie Beth
; APPLICANT: Siles-Santlago, Immaculada
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
; TITLE OF INVENTION: 12216, 17719, 41897, 47174, 33408, 10002, 16209, 314, 636,
; TITLE OF INVENTION: 27410, 33260, 619, 15985, 69132, 2158, 224, 615, 44373,
; TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55054, 16314, 1675, 9569 OR
; TITLE OF INVENTION: 13424 MOLECULES
; FILE REFERENCE: MP102-027P1MONMIM
; CURRENT APPLICATION NUMBER: US/11/312,958
; PRIOR FILING DATE: 2005-12-20
; PRIOR APPLICATION NUMBER: US/10/369,022
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: US 60/360,495
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/370,121
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/373,010
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/373,908
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/377,717
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US 60/379,949
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/382,409
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/385,280
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/386,879
; PRIOR FILING DATE: 2002-06-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 4234
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (863)...(2452)
US-11-312-958-25

Query Match      1.1%; Score 34; DB 7; Length 4234;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 70; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY      2240 TGATGGCATTCATCTTGTGTTCCGGAAGGAGATTCCTGCGGAGCCTCA 2299
DB      1908 TGTTCCGCACTTTCAGCTCTCCCGCACTCCAAAGGGGTCGAGATCTGGGCAAGACCT 1967
QY      2300 GCAAGCTCTCGGGGAGGGGCTCCAGAGCGATGTTCTTCCGTTCAACGAGTCTGCG 2359
DB      1968 TCGAGCCTTCATGAGAGGAGCTGGGCTGCTCATTTCTTCTTCTTCAATCGGGGATCC 2027
QY      2360 TATACAGCAG 2369
DB      2028 TCTTCTCCAG 2037

RESULT 13
US-10-505-928-409
; Sequence 409, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
```

/ PRIOR FILING DATE: 2002-03-07
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: PatentIn 3.2
/ SEQ ID NO 409
/ LENGTH: 4573
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-505-928-409

Query Match 1.1%; Score 34; DB 6; Length 4573;
Best Local Similarity 59.2%; Pred. No. 11;
Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1494 GTCTCCCAACTGAGCCCGACACCAAGCAGGCTCTCCCTTGATGAGCCCGCTGTGAA 1553
DB 437 GTCCGCGATTGGAGGACGAGAGCTCCAGAGTTCTCCCAACATCTCCAGAGCTGGA 496

QY 1554 TGACCAAGGCTGCCCCGAGCAGACGATGAGATGAGG 1591
DB 497 TTCCCGGCGCTGCACCTCGGAGAACGAGAAACGAGG 534

RESULT 14
US-10-505-928-47
/ Sequence 47, Application US/10505928
/ Publication No. US20060088532A1
/ GENERAL INFORMATION:
/ APPLICANT: Ludwig Institute for Cancer Research et al.
/ TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
/ FILE REFERENCE: 28967/39178
/ CURRENT APPLICATION NUMBER: US/10/505,928
/ CURRENT FILING DATE: 2004-08-27
/ PRIOR APPLICATION NUMBER: US 60/363,019
/ PRIOR FILING DATE: 2002-03-07
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: PatentIn 3.2
/ SEQ ID NO 47
/ LENGTH: 4585
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: KIAA0062
US-10-505-928-47

Query Match 1.1%; Score 34; DB 6; Length 4585;
Best Local Similarity 59.2%; Pred. No. 11;
Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1494 GTCTCCCAACTGAGCCCGACACCAAGCAGGCTCTCCCTTGATGAGCCCGCTGTGAA 1553
DB 449 GTCCGCGATTGGAGGACGAGAGCTCCAGAGTTCTGCCCAACATCTCCAGAGCTGGA 508

QY 1554 TGACCAAGGCTGCCCCGAGCAGACGATGAGATGAGG 1591
DB 509 TTCCCGGCGCTGCACCTCGGAGAACGAGAAACGAGG 546

RESULT 15
US-11-217-529-174434
/ Sequence 174434, Application US/11217529
/ Publication No. US20060099612A1
/ GENERAL INFORMATION:
/ APPLICANT: SUNTORY LIMITED
/ APPLICANT: NAKAO, YOSHIHIRO
/ APPLICANT: NAKAMURA, NORIHIISA
/ APPLICANT: KODAMA, YUKIKO
/ APPLICANT: FUJIMURA, TOMOKO
/ APPLICANT: ASHIKARI, TOSHIHIKO
/ TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
/ FILE REFERENCE: S-38-285
/ CURRENT APPLICATION NUMBER: US/11/217,529
/ CURRENT FILING DATE: 2005-09-02
/ PRIOR APPLICATION NUMBER: US 10/932,182

/ PRIOR FILING DATE: 2004-09-02
/ NUMBER OF SEQ ID NOS: 197023
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 174434
/ LENGTH: 555
/ TYPE: DNA
/ ORGANISM: Saccharomyces pastorianus
US-11-217-529-174434

Query Match 1.1%; Score 33.8; DB 7; Length 555;
Best Local Similarity 48.2%; Pred. No. 3.9;
Matches 95; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 2532 GGAGAAAGTGGATTGAGACATCAATGAGCCATTGACCTGGCGGAGAGAGCAGAGCCC 2591
DB 246 GACACAGCTGTTCTGACACAGAAAGCCATAGATCTGACAGAGAGCCCGACACAC 305

QY 2592 CCCCCCTGAGTTCTGGGCGAGAGCCCCCTGACACAAAGTCCCTGATGAAGCCACCGC 2651
DB 306 AAACACTACGATCAATTAATTAACAACATGACAGAGCCGCTTGAAGAGAGCCGA 365

QY 2652 GGCTGACCAAGATGACGAGATGACCTGAGGCGCTCGGCAATGCGTGAAGCGCAGGC 2711
DB 366 CAGCAGCATTAATTATTATCAAGAGAGAGGATCAATCAATTTCTGTGCGCTGATTTGCC 425

QY 2712 CCCGACCGCGGCAACA 2728
DB 426 CAGCAGCAGTACCAACA 442

Search completed: May 22, 2006, 11:25:34
Job time : 32.413 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 20, 2006, 01:28:11 ; Search time 14245 Seconds
(without alignments)
7036.685 Million cell updates/sec

Title: US-09-555-342B-2

Perfect score: 5463
Sequence: 1 MBEIQRPTGSRGLGAPENS.....SATSSARPHVLSHKESLVY 1045

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ .p2n.model -DEV=xlh
-Q=/abses/ABSSWEB.spool/US09555342/runat.19052006.121454.13184/app.query.fasta_1
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -HOST=abs07
-DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs07
-USRR=US09555342.OCEN.1.1.5548.0/runat.19052006.121454.13184 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_env:.*
2: gb_pat:.*
3: gb_ph:.*
4: gb_dl:.*
5: gb_pr:.*
6: gb_ro:.*
7: gb_sts:.*
8: gb_sy:.*
9: gb_un:.*
10: gb_vl:.*
11: gb_ov:.*
12: gb_hcg:.*
13: gb_in:.*
14: gb_om:.*
15: gb_ba:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5463	100.0	3442	2	CS247579 Sequence
2	5463	100.0	3442	5	AB008430 Homo sapi
3	5463	100.0	3546	5	BC041595 Homo sapi

4	5463	100.0	5028	2	CS130723	CS130723 Sequence
5	5459	99.9	4687	2	BD231200	Human cyt
6	5431.5	99.4	4932	5	BC071592	BC071592 Homo sapi
7	4095.5	75.0	2937	2	CQ724059	Sequence
8	3963.5	72.6	3902	6	BC030329	Mus muscu
9	2941.5	53.8	3997	2	CQ895396	Sequence
10	2941.5	53.8	3997	5	AB018336	AB018336 Homo sapi
11	2936.5	53.8	3997	2	CQ725044	Sequence
12	2924	53.5	3719	6	BC009153	BC009153 Mus muscu
13	2117	38.8	2632	6	BC004009	BC004009 Mus muscu
14	1879.5	34.4	4048	13	BT024186	BT024186 Drosophill
15	1784	32.7	2431	6	BC027077	BC027077 Mus muscu
16	1669	30.6	2596	5	BC021301	BC021301 Homo sapi
17	1539	28.2	2286	6	BC043327	BC043327 Mus muscu
18	1437	26.3	129380	11	AY739094	AY739094 Takifugu
19	1150.5	21.1	3200	2	CQ842377	CQ842377 Sequence
20	1150.5	21.1	3200	5	AK125336	AK125336 Homo sapi
21	1075	19.7	2436	11	BC110755	BC110755 Xenopus 1
22	913.5	16.7	1747	2	CQ575283	CQ575283 Sequence
23	905.5	16.6	1230	2	CQ572985	CQ572985 Sequence
24	905.5	16.6	2538	13	AY060484	AY060484 Drosophill
25	891.5	16.3	4150	2	CQ575282	CQ575282 Sequence
26	891.5	16.3	42014	12	AC018327	AC018327 Drosophill
27	891.5	16.3	160712	13	AC009537	AC009537 Drosophill
28	891.5	16.3	164712	13	AC007589	AC007589 Drosophill
29	891.5	16.3	303434	13	AB003604	AB003604 Drosophill
30	850.5	15.6	2888	4	BSM602608	BSM602608 Homo sapi
31	798.5	14.6	3474	4	BT008434	BT008434 Arabidops
32	798.5	14.6	3592	4	BT002491	BT002491 Arabidops
33	790	14.5	235700	12	AC130919	AC130919 Rattus no
34	790	14.5	283769	12	AC094777	AC094777 Rattus no
35	754.5	13.8	3275	2	CQ850384	CQ850384 Sequence
36	754.5	13.8	3275	5	AK127510	AK127510 Homo sapi
37	754	13.8	3594	11	CT025506	CT025506 Xenopus t
38	754	13.8	3760	11	BC096388	BC096388 Xenopus t
39	753	13.8	3488	11	BC060449	BC060449 Xenopus 1
40	752.5	13.8	3651	6	AB032828	AB032828 Rattus no
41	752	13.8	4051	6	AF152247	AF152247 Mus muscu
42	748.5	13.7	3309	2	CS032039	CS032039 Sequence
43	748.5	13.7	3309	2	CS032937	CS032937 Sequence
44	748.5	13.7	3309	2	CS040991	CS040991 Sequence
45	748.5	13.7	3309	2	CS041889	CS041889 Sequence

ALIGNMENTS

RESULT 1	CS247579	3442 bp	DNA	linear	PAT 09-JAN-2006
LOCUS	CS247579				
DEFINITION	Sequence 31 from Patent EP1612281.				
ACCESSION	CS247579				
VERSION	CS247579.1	GI:84661528			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Raponi, M.				
AUTHORS	Methods for assessing patients with acute myeloid leukemia				
TITLE	Patent: EP 1612281-A 31 04-JAN-2006;				
JOURNAL	Veridex, LLC (US)				
FEATURES	Location/Qualifiers				
source	1..3442				
ORIGIN	/organism="Homo sapiens"				
Alignment Scores:	/mol type="unassigned DNA"				
Score:	/db_xref="taxon:9606"				
	0	Length:	3442		
	5463.00	Matches:	1045		

Percent Similarity:	100.0%	Conservative:	0
Beet Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0
US-09-555-342B-2 (1-1045) x CS247579 (1-3442)			
QY	1	MetGlyGluIleGluGlnatgProthrProGlySerArgLeuGluValAProGluAenSer	20
DB	49	ATGGAGAAATATAGAGAGAGAGCCGACCCAGAGATCAAGACTGGGGGGCCCGGAAATTCG	108
QY	21	GlyTlIeserThrLeuGluuATgGlyGlnLysProProThrProSerGlyLysLeuVal	40
DB	109	GGGATCAATGACCTTGAAACGTGACAGAAACCCGCCCAACACCTTCAGAGAAATCTGTC	168
QY	41	SeTlIeLysIleGlnMetLeuAspAspThrGlnGluAaPheGluValAProGlnatrgAla	60
DB	169	TTCATCAAAATCCAGATGCTGAGATGACACCCAGAGAGCATTTGAAAGTTCCACAAAGACT	228
QY	61	ProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGluGlyAspTyr	80
DB	229	CTGGGAGAGGTGCTGGTGGATGCAAGTTTGCAACCACTCAACCTCGTGGAGGTGACTAT	288
QY	81	PheGlyLeuGluuPheProAspHisLysLysIleThrValTTrpleuAspLeuLysPro	100
DB	289	TTTGGCTCGAGTTTCTGTATCACAAAAGATCAAGGTGGCTGGATCTCTAAACC	348
QY	101	IleValIleGlnIleArgArgProLysHisValValLysPheValValLysPhePhe	120
DB	349	ATTGTGAACAGATTAGAGGCCAAAGCACTGTTGTTAAAGTTGTGGTGAATTCCTT	408
QY	121	ProProAspHisThrGlnLeuGlnGluLeuThrArgTyrLeuPheAlaLeuGlnVal	140
DB	409	CCGCTGACCAACACACAACTCCAAGAAAGACTCAAGAGTACCTGTTCCGGCTGCAGGTC	468
QY	141	LysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAlaAlaLeuLysIle	160
DB	469	AACACAGACTTGGCTCAAGGACAGTTGACGTGAATGACACAGCCAGCTCTTGATTT	528
QY	161	SeThrIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgGlnHisLeu	180
DB	529	TTCACATTTGTCMAATCTGAGATTTGGGGAATTTGATGAACTTGGACAGAGACCTTA	588
QY	181	AlaLysAsnLysTyrIleProGlnIleAspAlaLeuGluAspLysIleValGluPheHis	200
DB	589	GCAAAATATATATACATCTCAGCAAGACGCACTAGAGGACAAATTCGGAAATTTAC	648
QY	201	HisAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuGlnIleAlaArg	220
DB	649	CATTAACCACTTGACAAACACAGCAGATCAGATTTCCAGCTCTAGAGATTTGCCCT	708
QY	221	ArgLeuGlnMetTyrGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysIle	240
DB	709	CGGCTAGAGATGATGGAATCCGGTTGCCACCCGGCCAAAGACAGGAGGACAGAAATC	768
QY	241	AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAla	260
DB	769	AATCTGGCCCTTGCAACAGGGAATCTAGTGTTCACAGGGTTTCACTAAGATCAATGCC	828
QY	261	PheAsnThrAlaLysValArgLysLeuSerPheLysArgLysArgPheLeuIleLysLeu	280
DB	829	TTCAACTGGGCGCAAGGCGGAGCTGAGCTTCAAGAGGAAGGCTTTTCATCAAGCTC	888
QY	281	ArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluuPheLeuMetAlaSerArgAsp	300
DB	889	CGGCGAGATGCCAATAGTGGCTACAGAGATCACTTGAATTCCTGATGGCAATCCGGAT	948
QY	301	PheCysLysSerPheTyrLysIleCysValGluHisHisAlaPhePheArgLeuPheGlu	320
DB	949	TTCTGCAAGCTCTTCGAGAAATCTGTGTTGAAATCATCATCTCTTTAGACTTTTGA	1008
QY	321	GluProLysProLysProValLeuPheSerArgLysSerPheArgPheSer	340

DB	1009	GAGCCCAAAACAAAGCCCAAGCCCTCTCTTTAGCCGGGGGTCAATCTTCGTTCAAGT	1068
QY	341	GlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyHisLysLysValGln	360
DB	1069	GGTCCGACCTCAGAGAGAGTTCTCGACTATGTTAAAGAGAGACATTAAGAAAGTGCAG	1128
QY	361	PheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu	380
DB	1129	TTTGAAGAGAACACACAGCAAGATCTCACTTACCGGAGCTTGCTTCAACAGCTTACAGA	1188
QY	381	LeuAsnSerGluValLeuGlnGlnSerGlnGlnIleThrSerLeuThrPheGlyGlyGly	400
DB	1189	CTGAATTCGAAAGTGTGAGAGATCTTCAGCAGAGACCAAGCTTACATTTGGAGAAAGT	1248
QY	401	AlaGluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGly	420
DB	1249	GCGAATCTCCAGGGGGCCAGAGCTGCGCGAGAGAAAGAACAGAGTTTCCCGCGGG	1308
QY	421	GluProGlySerHisProSerProAlaProArgArgSerProAlaLysLysGlnAla	440
DB	1309	GAGCCGGGTCCACCCGAGCTGCGCGAGAGAGAGCCCGCGGTAAACAGCAGGCG	1368
QY	441	AspGlyAlaAlaSerAlaProThrGlnGlnGlnGlnValValLysAspArgThrGln	460
DB	1369	GAGGAGCCGCTCGGCGCCACGAGAGAGAGAGAGAGTGTGAATGAGATGACCAAG	1428
QY	461	GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeu	480
DB	1429	CAGAGTAACTCAGGCCCGCGAGCCCAAGCACAGGCTCTGATGCGGAGTCCCTACCTT	1488
QY	481	SerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro	500
DB	1489	TCCGAGCTCTCTGTGAATCCGAGGGGGAGTGGCCCTCCCAACGTGACTGTCTCC	1548
QY	501	AsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuLeuAsnArgIn	520
DB	1549	AACCTGAGCCCGACACCAAGAGGCTTCTCTTGATAGCCCGCTGTAAATGACCAAG	1608
QY	521	AlaCysProArgThrAspAspGluAspGluGlyArgArgLysArgPheProThrAspLys	540
DB	1609	GCCTGCCCGGACGAGATGAGAGTGAAGGCGCGAGAGAAAGATTTCCCACTGATAA	1668
QY	541	AlaTyrPheIleAlaLysGlnValSerThrThrGluArgThrTyrLeuLysAspLeuGlu	560
DB	1669	GCGTACTTATGCTAAAGAGATGCTTACACCGAGCGAACTATCTGAAGATTCGGA	1728
QY	561	ValIleThrSerTyrPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeu	580
DB	1729	GTTATCATCTTGCTGTTTACAGCACAGTACAGCAAGAGAGAGCCCATGCGGAGACTCG	1788
QY	581	LysSerLeuIlePheProAsnPheGlnProLeuHisLysPheHisThrAsnPheLeuLys	600
DB	1789	AAAAGTCTCATATCCCGAATTTTGAACCTTTGCACAATTTCACTAATTTTTCAGAG	1848
QY	601	GluIleGlnGlnArgLeuAlaLeuTyrGluGlyArgSerAsnAlaGlnIleArgAspTyr	620
DB	1849	GAATTTGACCAACGACTTGCCCTTGAGGAAGCCGCTCAAAATTCAGAGATTTAC	1908
QY	621	GlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaHisHis	640
DB	1909	CAGAATCGGAGATGCTACTGTAAGACATTCACGGCATTAAGCACTGCGGCTCAC	1968
QY	641	LeuTyrLysHisSerGlnAlaLeuGlnAlaLeuGluAsnGlyIleLysSerSerArgArg	660
DB	1969	CTGTGGAACACACAGGAGGCTTTGAGGGCCCTGGAAGATGGAATCAAGAGCTCCCGGG	2028
QY	661	LeuGluAsnPheCysArgAspPheGlnLeuGlnLysValCysTyrLeuProLeuAsnThr	680
DB	2029	CTGGAGAACTTCTGACAGACTTTGAGCTGCAGAAAGTGTTGTTACTTACCGCTCAACAC	2088
QY	681	PheLeuLeuArgProLeuHisArgLeuMetHisLysTyrLysGlnValLeuGluuArgLeuCys	700
DB	2089	TTCTCTCGGCGCACTGCAAGGCTCATGCACTACAGCAGGTCTTGAAGCGGCTGTGC	2148

QY 701 LyeHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIle 720
 DB 2149 AAcACACACCCGCGAGCCAGCCGACTTCAGGAGCGCGAGCGCTTGGCAGAAATC 2208
 QY 721 ThrGluMetValAlaGlnLeuHisGlyThrMetIleIleuMetGluAsnPheGlnIleu 740
 DB 2209 ACCGAGATGGGTGGCAGCTCCACGATACGATGATCAAGATGGAAATTTCCAGAACTG 2268
 QY 741 HisGluLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 760
 DB 2269 CACGAACTCAAGAAAGATTGATTGCAATTCATCTTGCTCCGGGAAGGAGATTC 2328
 QY 761 IleArgLeuGluSerLeuSerLeuSerGlyIleuGlnGlnIleuGlnIleuGlnIleu 780
 DB 2329 ATCCGTTGGGACCTTCAGACAGCTCTCGGGAAAGGGGCTCCAGCAGCGCATGCTTC 2388
 QY 781 LeuPheAsnAspValIleuLeuIleuIleuIleuIleuIleuIleuIleuIleuIleu 800
 DB 2389 CTGTTCAACGACGCTCGCTATACAGAGCGGGGGCTGACGGCTCCATTCAGTTTAA 2448
 QY 801 ValHisGlyGlnLeuProLeuIleuIleuIleuIleuIleuIleuIleuIleuIleu 820
 DB 2449 GTCACACGGGAGCTCCGCTCTATAGCATGAGATTGAGAGAGCGAAGACGAGTGGGG 2508
 QY 821 ValProHisGlyLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 840
 DB 2509 GTGCGCCACGCTGCTGACCTCCGGGGCCAGGGGCGCATCGCTCATCGGCGCCGATCT 2568
 QY 841 ArgSerGluMetGluIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 860
 DB 2569 CGGTCGAGATGGAGAAATGGGTTGAGAGACATCCAGATGGCCATTCCTGGCGAGAG 2628
 QY 861 SerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnIleuSerProAsp 880
 DB 2629 AGCAGACGCCCCCGCTGATCTGCGCCAGCGCCCCCTGCAACAAGTCCCTGAT 2688
 QY 881 GluAlaLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 900
 DB 2689 GAAAGCCACCGCGGTGACAGAGATCAAGAGATCACTGAGCGCTTGGCGACATCCGCTG 2748
 QY 901 GluArgGlnAlaProHisArgGlyAsnThrMetValHisValCysTrpHisArgAsnThr 920
 DB 2749 GAGCGCCAGGCGCCGACCGCGGCAACATGGTGCACGCTGCTGGCACCGCAACACC 2808
 QY 921 SerValSerMetValAspPheSerIleAlaValGluAsnGlnIleuSerGlyAsnIleu 940
 DB 2809 AGCGTCTCCATGGGACTTCAGCATCGCAGTGGAGAAATCAGTTGCTCGAAACCTGCTG 2868
 QY 941 ArgIleuPheIleuAsnSerAsnGlyTrpGlnIleuIleuIleuIleuIleuIleuIleu 960
 DB 2869 AGAATAATCAAAACACAAACGGGTGGCAGAAAGTGGGGGTGGTTCACAAACTTCTGC 2928
 QY 961 LeuPhePheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 980
 DB 2929 CTGTTCTTCAAAATACACACAGCAATCATCCCTTGGCAGCTGCTGCTGGCC 2988
 QY 981 TyrSerLeuThrIleProSerGluSerGluAsnIleGlnIleuAspIleuIleuIleu 1000
 DB 2989 TACTCGTCAACATCCCTCTGAGTCCGAGAAACATCCAGAAAGACTGCTGCTCAACTG 3048
 QY 1001 HisPheIleuSerHisValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 1020
 DB 3049 CACTTCAGTCCCAACGCTACTACTTCAAGGGCGGAAAGCAGATGACAGTTGCAAGGTGG 3108
 QY 1021 MetGluValIleuArgSerAlaThrSerSerAlaSerArgProHisValIleuSerHisLys 1040
 DB 3109 ATGGAAGTGAATCCGCAATGCGACAGCTCTGCTCGGAGCCCAACGCTTGAAGCCAAA 3168
 QY 1041 GluSerLeuValIleu 1045
 DB 3169 GAGTCTCTGTGTAT 3183

RESULT 2
 AB008430
 LOCUS Homo sapiens mRNA for CDEP, complete cds.
 DEFINITION
 ACCESSION AB008430
 VERSION AB008430.1 GI:2766164
 KEYWORDS CDEP.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS Koyano, Y., Kawamoto, T., Shen, M., Yan, W., Noshiro, M., Fujii, K. and Kato, Y.
 TITLE Molecular cloning and characterization of CDEP, a novel human protein containing the ezrin-like domain of the band 4.1 superfamily and the Dbl homology domain of Rho guanine nucleotide exchange factors
 JOURNAL Biochem. Biophys. Res. Commun. 241 (2), 369-375 (1997)
 PUBMED 9425278
 REFERENCE 2 (bases 1 to 3442)
 AUTHORS Koyano, Y., Kawamoto, T. and Kato, Y.
 DIRECT SUBMISSION
 JOURNAL Submitted (22-OCT-1997) Takeshi Kawamoto, Hiroshima University School of Dentistry, Department of Biochemistry, 1-2-3 Kasumi Minami-ku, Hiroshima, Hiroshima 734, Japan (E-mail: k Kawamoto@ipc.hiroshima-u.ac.jp, Tel: 082-257-5688, Fax: 082-257-5629)
 FEATURES
 source
 1..3442
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_type="Chondrocyte"
 /tissue_type="cartilage"
 /dev_stage="embryo"
 49..3186
 /function="Rho Guanine Nucleotide Exchange Factor"
 /note="Band 4.1 superfamily"
 /codon_start=1
 /product="CDEP"
 /protein_id="BA24267.1"
 /db_xref="GI:2766165"
 /translation="MSGIEKRPPTGSRRLAPENSGISTLERGCKPPTPSGKIVSIKI QMLDQEAPEVFORAPGVLDVACNHLNVEGDFGLFPHKKTIVLDLKPVI KOIRPKHVVVKKVVEFPEDHTLOEBELRYLFAIQVQDLOGLTCDNDSALLI SHIVQEGIDFEDLREHLAKNKYIPODDALEDKIVFEHNNHIGQTPAESDQLLEI ARLEMYGRIHPAKOREGKINLAVANTGILVFOGFTKINAFNAKVRLLSPKRRF LTKLRDANSAYODTLEPIMASRDPCKSPFKICVEHHAAPRLFEEDKPKPKPVLFSG SSFRSGARKOVLDYKXGSGHKQVPERKSKIHISIRLSAQPTELNSVLEDSQS TSLTFEGBASPEGQSCKRGKEPKVAGSGSPSPAPRSPKGNQADDAASAPTEE EEEVVNDRTQOSKPQPPSTGSLTGSPHLSELVSVSOGVAPANTVLSNLSPDTKO ASPLIPLLDQACPRPTDDEBGRKRPFTDKAYFLAKEVSTERVYLKDLVITSMF OSTVSKEDMPKALSLIFPNPEPLKHPHTNFKETLEORLALWEGSNQIDRYORIG DWLKVIOGMKTHAHLMAKRSALTELENGICYSRLEMLNCPPELOKVCYVLIANFTL LRPRLHMYKQVLERLCKHNPSSHADFPDCCRALALETBMVAQLGTMKXENFOUL HEIKQULIGDNLVVGREFIRLGSLSKSGKLGQRMFLFNDVLLVTSRGLTASID FKVHGQPLPYGMITIESBDEWGVHCLITLRGQSLIIVASRSEMEKWEEDIQMAD LKESSSPAPEPLASPPPNKSPDEATLADQESBDDLSASTLEREOAPRGNTMNV CHMRNTSVSNVPLASIVENOLSGNLTKRFNSNGWQKLVVFTNFCLEPKSHQDNHP LSLPLIGSLTIPSESENIQKDYVVKLHPKSHVYFRASEVTTERMMEVIRSATSS ASRPVLSHKESLVY"
 3442
 /note="50 a nucleotides"

polya_site
 ORIGIN
 Alignment Scores:
 Pred. No.: 0
 Score: 5463.00
 Percent Similarity: 100.0%
 Best Local Similarity: 100.0%
 Query Match: 100.0%
 Length: 3442
 Matches: 1045
 Conservative: 0
 Mismatches: 0
 Indels: 0

DB: 5 Gaps: 0
US-09-555-342B-2 (1-1045) x AB008430 (1-3442)
QY 1 MetGlyGluIleGluGlnArgProThrProGlySerArgLeuGluValAProGluAAsnSer 20
DB 49 ATGGGAGAAATAGAGAGAGGCGACCCCGAGATCAAGACTGGGGGCCCCGGGAAATTGCG 108
QY 21 GlyTLeSerThrLeuGluValArgGlyGluValAspProProThrProSerGlyValLeuVal 40
DB 109 GGGATACGATCTTGGAACGTGACAGAACCGCCCCCAACACCTTCAGGAAAACTCGTG 168
QY 41 SerTLeValIleGlnMetLeuAspPheThrGlnGluValAAspGluValAProGlnArgAla 60
DB 169 TCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAGTTCACAAAGAGCT 228
QY 61 ProGlyValValLeuLeuAspAlaValCyAspAspHisLeuAsnLeuValGluGlyAspTyr 80
DB 229 CCGGGAGAGGTGCTGTGATGATGACACCCAGAGCATTTGAGTTCACAAAGAGCTAT 288
QY 81 PheGlyLeuGluPheProAspHisValValIleThrValTLeuAspLeuLeuValPro 100
DB 289 TTGGCTCGAGTTCTCTGATCAAAAAGATCAAGGTGTGGATCTCTTAAACCC 348
QY 101 ILeValIleGlnIleArgArgProLysHisValValIleValPheValValIleValPhePhe 120
DB 349 ATTGTAAACAGATTGGAAGGCCAAAGCACGTTGTTTAAAGTTTGTGGTGAATTTCTTT 408
QY 121 ProProAspHisThrGlnLeuGlnGluGluLeuThrArgTyrLeuPheAlaLeuGlnVal 140
DB 409 CCGCTCGACACACACAACTCCAAAGAAAGACTCAAGAGTACCTGTTCGGCTGACAGGTG 468
QY 141 LysGluAspLeuValGlnGlyArgLeuThrCyAspAspPheSerAlaIleLeuLeuIle 160
DB 469 AAGCAGAGACTTGCTCAAGCAGAGTTGACGTGTAAGACACACGCGAGCTCTTGGATT 528
QY 161 SerHisIleValGlnSerGluIleGlyAspPheAspGluValLeuAspArgGluIleLeu 180
DB 529 TCACATGTTGTGCAATCTGAGATTGGGGATTGTTGATGAAGCTTTGACAGAGACCTTA 588
QY 181 AlaLysAsnLysTyrTleProGlnGlnAspAlaLeuGluAspLysIleValGluPheHis 200
DB 589 GCAGAAAATTAATATACACTCTCAGCAAGACGCACTGAGAGCAAAAATCCGGGAATTTCC 648
QY 201 HisAspHisIleGluGlnThrProAlaGluSerAspPheGlnLeuGluIleAlaArg 220
DB 649 CATACCACTTTGACAAACACACAGCAAGATCCAGATTTCCAGCTCTAGAGATTGCCCT 708
QY 221 ArgLeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGluIleTyrHisIle 240
DB 709 CGGCTAGAGATGTATGGAATCCGGTTGCACCCGGCCAAAGSACAGGAGGACAGAGATC 768
QY 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAla 260
DB 769 AATCTGGCCGTTGCCAAACAGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC 828
QY 261 PheAsnThrAlaLysValArgLysLeuSerPheLysArgTyrAspPheLeuIleLysLeu 280
DB 829 TTCAACTGGGCCAAGGTGCGAAGCTGAGCTTCAAGAGAAAGGCTTTCTCATCAAGCTC 888
QY 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAlaSerArgAsp 300
DB 889 CGGCGAGATGCCAATAGTGTGACAGAGATACCTTGAATTTCTGATGGCCAGTCCGGAT 948
QY 301 PheCysLysSerPheThrLysIleCyValGluHisAlaPhePheArgLeuPheGlu 320
DB 949 TTCTGCAAGCTCTTCGAAATCTGTGTGAACATCATCTCTTTAACTTTTGA 1008
QY 321 GluPheLysProLysProValLeuPheSerArgLysSerPheArgPheSer 340
DB 1009 GAGCCCAACCAAGGCCAGGCCGCTCTTTAGCCGGGGGATCATCTTTCCGCTCAGT 1068
QY 341 GlyArgThrGlnLysGlnValLeuAspTyrValIleGluGluGlyHisValGln 360

DB 1069 GGTCCGAGCTCAGAACAGGTTCTCAGATGTAAAGAGAGACATAGAGAGGTGCAG 1128
QY 361 PheGluArgLysHisSerTyrIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
DB 1129 TTGGAAGGAAACACAGCAAGATTCAATTCTATCCGAGCCTTGCTTACAGCTTACAGAA 1188
QY 381 LeuAsnSerGluValLeuGlnGlnSerGlnGlnSerThrSerLeuThrPheGluGluGly 400
DB 1189 CTGAATTCGAAGTGTGGAGCAGTCTCAGCAGAGACACAGCTTATCTTGGAGAAAGGT 1248
QY 401 AlaGluSerProGlyGlyGlnSerCyValArgGlyLysGluProLysValSerAlaGly 420
DB 1249 GCCGAATCTCCAGGGGGCCAGAGCTGCCGCCAGGAAAGAAACCGAAGGTTTCCGGGG 1308
QY 421 GluProGlySerHisProSerProAlaProArgSerProAlaGlyAsnLysGlnAla 440
DB 1309 GAGCCGGGGTCCGACCCGAGCCCTGGCCGAGGAGAAAGCCCGGGGTAAACAGCAGGCG 1368
QY 441 AspGlyAlaAlaSerAlaProThrGluGluGlnGluValValLysAspArgThrGln 460
DB 1369 GAGGAGCGGCTCGGCGCCACGAGAGAAAGAGAGAGTGTTAAGATAGGACCCAG 1428
QY 461 GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeu 480
DB 1429 CAGAGTAAACTCAGCCCGCCAGCCACAGCAGGCTCCTCAGCTGGCAGTCTCACCTT 1488
QY 481 SerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro 500
DB 1489 TCCGAGCTGTGTGAATCTGCAAGGGGAGAGGCCCTTCGCAACGTGACTGTGTCTCC 1548
QY 501 AsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuAsnAspGln 520
DB 1549 AAGCTGAGCCCGGACACCAAGAGGCTCTCCCTTGATCAGCCCGCTGTGAATGACAG 1608
QY 521 AlaCyProArgThrAspAspGluAspGluGlyArgLysPheProThrAspLys 540
DB 1609 GCCTGCCCCGAGCAGATGAGTGAAGGCGCGAGGAAAGATTTCCCACTGTAATA 1668
QY 541 AlaTyrPheIleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAspLeuGlu 560
DB 1669 GGATCTTATAGCTTAAGAGAGTGTCTCACCGAGCAATATTTGAAGATCTCGAA 1728
QY 561 ValIleThrSerTyrPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeu 580
DB 1729 GTTATCACTTCGTTGTTTCAAGACAGTGAAGCAAGAGCCATGCGGAAAGCAGCTG 1788
QY 581 LysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAsnPheLeuVal 600
DB 1789 AAAAGTCTATATTCGGAATTTTGAACTTTGCACAAATTTCAATATTTTCTCAAG 1848
QY 601 GlnIleGluGlnArgLeuAlaLeuTyrGluArgSerAsnAlaGlnIleArgAspTyr 620
DB 1849 GAATTTGACCAAGCACTTCCCTGTGGAAAGCCGCTCAAAATGCCAAATCAGAGATTAC 1908
QY 621 GlnArgIleGlyAspValMetLeuLysAsnIleGlnIleMetLysIleLeuAlaAlaHis 640
DB 1909 CAAAGATCGGGAGATGTATGTGAAGAACTTACAGGCGCATGAAGCACTGGGGCTCAC 1968
QY 641 LeuThrLysHisSerGluAlaLeuGluValAlaLeuGluAsnGlyTleLysSerSerArgArg 660
DB 1969 CTGTGGAAACACAGCAGGCGCTTGGAGGCTCGGAAATGGAAATCAAGAGCTCCGGCG 2028
QY 661 LeuGluAsnPheCyValArgAspPheGluGluGlnLysValCyValTyrLeuProLeuAsnThr 680
DB 2029 CTGGAGAACTTCTGAGAGACTTTGAGCTGCGAAGGTGTGTAACCTACCGCTCAACACC 2088
QY 681 PheLeuLeuArgProLeuHisArgLeuMetHisTyrLysGlnValLeuGluArgLeuCyVal 700
DB 2089 TTCTCTCTGCGGCACTGACCGGCTCATGCACTACAAAGAGTCTTGGAGCGGCTGTGC 2148
QY 701 LysHisHisProProSerHisAlaAspPheArgAspCyValArgAlaAlaLeuAlaGluIle 720

D	2149	AAACACACCCGCGAGCCAGCCGACTTCAGGAGCTCCGAGCGCTTGAGCAGAGATC	2208
Q	721	ThGtGwctVAlAaInLeuHlglYThrMetIlyvMetGluAaPPhgInlybLeu	740
D	2209	ACGAGAGATGGGACAGCTCCACGGTACGATGACAAAGATGAGAAATTCACAGAGCTG	2268
Q	741	HlglLeuLeuLybLybAaPLeuIleGlYlLeaPbaNleuValProGlYArgGluphe	760
D	2269	CACGAACTCAAGAAAGATTTATGATGATGCAATCTTGCTGTCGGGAAAGGAGATTC	2328
Q	761	lLeArgLeuGlYSerLeuSerLybLeuSerGlYlybGlYleuGlnGlnArgMetPhe	780
D	2329	ATCCGTCTGGGACAGCTCAGCAAGCTCTCGGGAAGGGGCTCCAGCAGCATGTTCTTC	2388
Q	781	LeuPheAaAaPValLeuLeuYTrThrSerArgLybLeuThrAlaSerAaGlnPheLyb	800
D	2389	CTGTTCACAGACGCTCTGCTATACAGAGCCGGGGGCGAGCGGCTCCAAATCAGTTTAA	2448
Q	801	ValHlglYlInLeuProLeuYTrGlYMetThrIleGlYlInLeuSerGlYUaPglY	820
D	2449	GTCACGGGAGCTCCGCTCTATGAGATGACGATGAGAGAGAGAGAGAGAGAGAGAG	2508
Q	821	ValProHlscYbLeuThrLeuArgGlYlInArgInSerIleIleValAlaSerSer	840
D	2509	GTCGCCACCTGCTGACCTCCGGGGCAGCGGCGAGCTCATGCTGCGCCGAGTTCT	2568
Q	841	ArgSerGlMetGlYbVTrValGlYUaPglIleGlYMetAlaIleAaPLeuHlglYb	860
D	2569	CGGTCCAGATGAGAAATGGGTGAGAGATCCAGATGGCCATTGACCTGCGGAGAG	2628
Q	861	SerSerSerProAlaProGlYbPheLeuAlaSerSerProPheAaPLeuLybSerProAaP	880
D	2629	AGCAGACAGCCCCCGCTGATCTCTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2688
Q	881	GluAlaThrAlaAlaAaPglInLeuSerGlYUaPbaPLeuSerAlaSerArgYThrSerLeu	900
D	2689	GAAACACACCGGGCTGACAGAGATCAAGATGATGACGCTCGCGACATCGCTG	2748
Q	901	GlYUaPglInAlaProHlscYbLeuArgLybAaPLeuValHlglYbVTrPheAlaPLeuThr	920
D	2749	GAGGCGCAGGCGCGGACCGGCAACCAATGAGTGGCTGCTGGCACCGCAACCC	2808
Q	921	SerValSerMetValAaPheSerIleAlaValGlYUaPglInLeuSerGlYUaPLeuLeu	940
D	2809	AGCCTCTCCATGATGAGATTCAGCATTCAGATGAGATTCAGTGTCTCGGAAACCTGCTG	2868
Q	941	ArgLybPheLybAaPLeuSerAaPglYTrPglInLybLeuTrValAlaPheThrAaPheCys	960
D	2869	AGGAAATTCAAAACAGCAACGGGTGGCAGAAAGCTGGGTGGTTCACAAACTTCTGC	2928
Q	961	LeuPhePheYTrLybSerHlglInAaPbaHlscPheLeuAlaSerLeuProLeuLeuGlY	980
D	2929	CTGTCTCTTCAAAATCAGACAGAGACAAATATCCCTTGGCACCTGCTCTCTCGGC	2988
Q	981	YTrSerLeuThrIleProSerGlYUaPglInLeuIleGlYbAaPglYValPheLybLeu	1000
D	2989	TACTCGCTCAACATCCCTCTGAGTCCGAGAAACATCCGAAAGACTAGCTTTCAAGCTG	3048
Q	1001	HlglPheLybSerHlscYbValYTrPheArgAlaGlYUaPglYTrPheGlYUaPglY	1020
D	3049	CACTTCAAGTCCACGCTCTACTTACGGGCGGAAACGAGTACAGTTGGAAGGTGG	3108
Q	1021	MetGlYUaPglInLeuArgSerAlaThrSerSerAlaSerArgProHlscValLeuSerHlscLyb	1040
D	3109	ATGGAAGATCCGACAGTGCACAGCTCTGCTCGACCCACGTTGTGAGCCACAA	3168
Q	1041	GlYUaPLeuValYTr 1045	
D	3169	GAGTCTCTGTGTAT 3183	
RESULT 3			
BC041595			
LOCUS	BC041595	3546 bp	mRNA linear PRI 08-MAR-2005

DEFINITION	Homo sapiens FER1, RhoGEF (ARHGEF) and plectrokin domain protein 1 (Chondrocyte-derived), transcript variant 1, mRNA (cDNA clone MGC:52223 IMAGE:5443230), complete cds.
ACCESSION	BC041595
VERSION	BC041595.1
KEYWORDS	GI:27370806
SOURCE	MGC.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 3546) Strauberg R, Feingold BA, Grouse LH, Derge JG, Klausner RD, Collins PS, Wagner L, Shennan CM, Schuler GP, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max ST, Wang J, Hsieh F, Diachenko L, Maruina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Udelin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullaby SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JB, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Kettelman M, Madan A, Rodriguez S, Sanchez A, Whitting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW, Touchman JW, Green ED, Dickson WC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butlerfield YS, Krzywinski MI, Skalska U, Smalins DE, Scherch A, Schein JE, Jones SJ and Marra MA. Mammalian Gene Collection Program Team Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
CONSRPT	2 (bases 1 to 3546) Director MGC Project.
JOURNAL	Direct Submission Submitted (20-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgabbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nhi.nih.gov Ahter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, O.L., Masello, C., Maskeri, B., Mastriani, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Wehrby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov Series: IRAL Plate: 44 Row: 1 Column: 24. Location/Qualifiers 1..3546 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:52223 IMAGE:5443230" /tissue_type="Brain, astrocytoma, grade IV" /clone_lib="NIH_MGC_98" /lab_host="DH10B" /note="Vector: pOTB7" 1..3546
gene	

/gene="FARP1"
/note="synonyms: CDEP, MG87400"
/db_xref="GeneID:10160"
/db_xref="MIM:602654"
266..3403
/gene="FARP1"
/codon_start=1
/product="FERM, RhodGF, and pleckstrin domain protein 1, isoform 1"
/protein_id="AAH41595.1"
/db_xref="GI:49898405"
/db_xref="GeneID:10160"
/db_xref="MIM:602654"
/translation="MGSEIQRPPTGSRNLGAPENSGISTLERGQKPPETWSGLVSIKI
OMLDDQAEVFPQAPRGKVLDAVCNHLNLYVEGYGLEFPFHKKITVWLDLKPLV
KQIRRPKHVVVKRVKPRPDHLOLELTRYLPALOVQODLQOGLTQNDTSALLI
SHIVQSEIDGFDALDREHLAKKTYIQODALDKIVFHHNTGOTPAESDQLLEI
ARLEWYGLRLHAKDREGTKINLAVANTGILVQGTGKINAFNAKVRSLSPKRRP
LILRPDANSAYODTLEFLMASRDFCKSPFKICVFNHAPFRLPEEPKPYLFRSG
SFRFSGRTQKQVLDVYEGGHHKVOGERGSKHSIRSLASPTLEINSEVLEQOOS
TSLTPGSGASPGGSGCRGKEPKVSGASHPSPAPRSPAGNQAODAPATPER
BEEVVDRTQASRQPPQPSGTSLSPHSELNVNSQGVAPANTLSPNLSPPDKO
ASPLISPLNDQACPTDDEBGRKRFPYDKAVYIAKEVSTTERTYLKDLEVTISWF
QSTVSKEDAMPKALKSLIFENFEPLKPHNPLKEIEQRLALWEGSNMADIRYQIG
DVLAKNIQGMKGLAAHLWKISEALEALENGIKSRLEPCRDPELQKCYPLANTFL
LRHLHMYKQVLERLCKHPPSHADPRDRAALAEITMVNQLHGTMIKMNPKOL
HELKDLIGIDNLVPGREFIRLGSLSKLSGKLQOQMPFLFNDVLLYTSRGLTASNO
FKHKGQPLVGMWTEIEESSEDEKWPCHCLTLRGQOSTIIVASSSEKWEVEDIQMAD
LAKSSPAPEFLASSPDKNSPDEATPAQDSEDDLSARISLEKQAPKSHQTHV
CMHRNTSVSWDPSIAVENQLSGNLRKFNKNSGMQKLWVFTNPLCLFPKXKQNDHP
LASPLPLGSLTIIPSESENTOKDYVFLHFKSHVYFRASEYTFERMEVIRSAVSS
ASRPVLSHRESLVY"

US-09-555-342B-2 (1-1045) x BC041595 (1-3546)
QY 1 MGGLVGLIUILEGIUGINARGPProthPrGGLYSerArgLeuGLIUAIProGLIUAenSer 20
DB 266 ATGGAGAAATAGAGCAGAGCCGACCCGAGATCACGACTGGGGGCCCGGAAAAATTCG 325
QY 21 GLYIleSerThrLeuGLIUAArgGLIULYsProProProThrPProSerGLYULYsLeuVal 40
DB 326 GGGATCAGTACCTTGGAACTGGACAGAACCGGCCCAACCTTCAGAAAAATCGTG 385
QY 41 SerIleLYsIleGIInMetLeuAspAspThrGInGLIUAIPheGLIULVALProGInAAGALA 60
DB 386 TCCATCAAAATCCAGATGCTGAGTGAACACCCAGAGAGCATTTGAAAGTTCCACAAAGGCT 445
QY 61 ProGLIULYsValIleuLeuAspAlaValCYAsnHisIleuAsnLeuValGLIULYsAspTYr 80
DB 446 CCTGGGAAGGTGCTGCTGGATGCGAGTTTGCAACCACTCAACCTCGTGGAGAGTGACTAT 505
QY 81 PheGLIULeuGLIuPheProAspHisIleLYsValIleThrValTTrPLeuAspLeuLeuLYsPro 100
DB 506 TTTGGCCTGAGTTTCTGTATCACAAAAAATGATCACGGTGTGGCTGATCTCTRAAAACC 565
QY 101 IleValIleGInIleArgArgProLYsHisValValIYsPheValValLYsPhePhe 120
DB 566 ATGTGTAACAGATTGAAAGCCCAAGCAAGCTGTGTCTTAAGTTTGTGTGAATTCCTTT 625
QY 121 ProProAspHisThrGInLeuGLInGLIuLeuThrArgTYrIleuPheAlaLeuGInVal 140
DB 626 CCGCCTGACCAACAACCTCCAAAGAACTCAAAAGTAACCTGTGTGGCTGCGAGGTTG 685
QY 141 LyGLInAspLeuAlaGInGLIArgLeuThrCYAsnAspThrSerAlaAlaLeuLeuIle 160

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3546
Score: 5463.00 Matches: 1045
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0

US-09-555-342B-2 (1-1045) x BC041595 (1-3546)
QY 1 MGGLVGLIUILEGIUGINARGPProthPrGGLYSerArgLeuGLIUAIProGLIUAenSer 20
DB 266 ATGGAGAAATAGAGCAGAGCCGACCCGAGATCACGACTGGGGGCCCGGAAAAATTCG 325
QY 21 GLYIleSerThrLeuGLIUAArgGLIULYsProProProThrPProSerGLYULYsLeuVal 40
DB 326 GGGATCAGTACCTTGGAACTGGACAGAACCGGCCCAACCTTCAGAAAAATCGTG 385
QY 41 SerIleLYsIleGIInMetLeuAspAspThrGInGLIUAIPheGLIULVALProGInAAGALA 60
DB 386 TCCATCAAAATCCAGATGCTGAGTGAACACCCAGAGAGCATTTGAAAGTTCCACAAAGGCT 445
QY 61 ProGLIULYsValIleuLeuAspAlaValCYAsnHisIleuAsnLeuValGLIULYsAspTYr 80
DB 446 CCTGGGAAGGTGCTGCTGGATGCGAGTTTGCAACCACTCAACCTCGTGGAGAGTGACTAT 505
QY 81 PheGLIULeuGLIuPheProAspHisIleLYsValIleThrValTTrPLeuAspLeuLeuLYsPro 100
DB 506 TTTGGCCTGAGTTTCTGTATCACAAAAAATGATCACGGTGTGGCTGATCTCTRAAAACC 565
QY 101 IleValIleGInIleArgArgProLYsHisValValIYsPheValValLYsPhePhe 120
DB 566 ATGTGTAACAGATTGAAAGCCCAAGCAAGCTGTGTCTTAAGTTTGTGTGAATTCCTTT 625
QY 121 ProProAspHisThrGInLeuGLInGLIuLeuThrArgTYrIleuPheAlaLeuGInVal 140
DB 626 CCGCCTGACCAACAACCTCCAAAGAACTCAAAAGTAACCTGTGTGGCTGCGAGGTTG 685
QY 141 LyGLInAspLeuAlaGInGLIArgLeuThrCYAsnAspThrSerAlaAlaLeuLeuIle 160

DB 686 AAGCAGACCTTGCGCTCAAGCAGAGTTGACGTGTAATGACACCGCAGCTCTTGATTT 745
QY 161 SerHisIleValGInSerGInIleGIYAspPheAspGLIUALeueuAspArgGLIULYsLeu 180
DB 746 TCACACATTTGTCAATCTAGATTTGGGATTTTGAATGAAGCTTTGACAGAGCACTTA 805
QY 181 AlaLYsAsnLYsTYrIleProGInGInAspAlaLeuGLIuAspLYsIleValGluPheHis 200
DB 806 GCAAAAATTAATTAATCACTCACTCAAGACGCACTAGAGGACAAAATTCGTGAATTTCA 865
QY 201 HisAsnHisIleGIYGLInThrProAlaGLIuSerAspPheGLInLeuGLIULYsAlaArg 220
DB 866 CATTAACCAATTCGAAACAAACACACGACGATCATGATTTCCAGCTCTTACGATATGCCCT 925
QY 221 ArgLeuGLInMetTYrGInIleArgLeuHisProAlaLYsAspArgGLIULYsThrLYsIle 240
DB 926 CGGCTAGAGATGTATGAAATCCGTTTGCAACCCGACAGGACAGGAGGACGAAAGATC 985
QY 241 AsnLeuAlaValAlaAsnThrGLYIleLeuValPheGLInGLYpHeThrLYsIleAsnAla 260
DB 986 AATCTGCGCTTGCCCAACACGGGAATTCATGTTTCAAGGTTTCACTAAGATCAATGCC 1045
QY 261 PheAsnTTrAlaLYsValArgLYsLeuSerPheLYsArgLYsArgPheLeuIleLYsLeu 280
DB 1046 TTCAACTGGGCCCAAGGTGCGAAAGCTGAGACTTCAAGAGAAAGCGCTTTCATCAAGCTC 1105
QY 281 ArgProAspAlaAsnSerAlaTYrGInAspTTrLeuGLIuPheLeuMetAlaSerArgAsp 300
DB 1106 CGGCCAGATGCCCAATGTGTCGACGAGATCACTTGAAATTCCTGATGGCCAGATCGGGAT 1165
QY 301 PheCYeLYsSerPheTTrLYsIleCYeValGInHisAlaPhePheArgLeuPheGLI 320
DB 1166 TTCTGCAAGTCTTCTCGAAAAATCTGTGTGAACATCAATGCTTCTTGAGACTTTTGA 1225
QY 321 GluProLYsProLYsProLYsProVALLeuPheSerArgGLYSerSerPheArgPheSer 340
DB 1226 GAGCCCAACCAAAAGCCCAAGCCCGTCTTATGACCGGGGTCATCATTTCCGTTCAAGT 1285
QY 341 GLYArgThrGInLYsGInValIleuAspTYrValIYsGLIULYsIleLYsLYsValGIn 360
DB 1286 GGTCCGACCTCAGAAAGAGAGGTTCGACTGATGTTAAAGAAAGAGACATTAAGAAAGTGAC 1345
QY 361 PheGLIULYsHisSerLYsIleHisSerIleArgSerLeuAlaSerGInProThrGIn 380
DB 1346 TTTGAAGAAGACACAGCAAGATTCATTCATCCGAGCTTCTTCAACGCTTACAGAA 1405
QY 381 LeuAsnSerGLIUALeugInSerGInSerThrSerLeuThrPheGLIULYs 400
DB 1406 CTGAATTCGAAAGTCTGAGAGCTCTCAGACAGACACCACTTCAATTTGAGAAAGGT 1465
QY 401 AlaGLIuSerProGLIYGLIuInSerCYsArgArgGLYLYsGLIuProLYsValSerAlaGLY 420
DB 1466 GCCGAATTCACAGGGGGCCAGAGCTGCGCGAGGAAAGAAACCAAGGTTTCCGCGGG 1525
QY 421 GluProGLYSerHisProSerProAlaProArgArgSerProAlaLYsAsnLYsGInAla 440
DB 1526 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTAAACAAGCGCG 1585
QY 441 AspGLYAlaAlaSerAlaProThrGInGLIuGLIuValIYsAspArgThrGIn 460
DB 1586 GACGAGACCGCTCGCGCCCAAGGAGAAAGAGAGAGTCTGTTAAGGTAGAACCAAG 1645
QY 461 GInSerLYsProGInProProGInProSerThrGInSerLeuThrGLYSerProHisLeu 480
DB 1646 CAGAGTAACTTCAGGCCCCCGAGCCCAACACAGCGCTCCCTGACTGACCTTCACCTT 1705
QY 481 SerGLIuSerValAsnSerGInGLYValAlaProAlaAsnValThrLeuSerPro 500
DB 1706 TCCGAGCTGTCTGTAACTTCGACGGGGAGAGTGCCCTTGCACACTGACCTGTCTCCC 1765
QY 501 AsnLeuSerProAspThrLYsGInAlaSerProLeuIleSerProLeuLeuAspGIn 520

Dp	1766	AACTGACGCCCCGACACCAAGCAGGCGCTTCCCTTGATCAGCCGCGCTGTAATGACCAAG	18235
Qy	521	AlAcysProAlyThrAspAspGluAspGluGlyAspArgLysArgPheProThrAspLys	540
Dp	1826	GCCTGCCCCCGGAGCGAGCATGAGGATAGAGGGCCGGAGAGAGATTTCCCAACTGATGAA	18855
Qy	541	AlArgPheIleAlaLysGluValSerThrThrgluValgThrTYrLeuLysAspLeuGlu	560
Dp	1886	GCAGTACTTCATAGCTAAGAGAAAGTGCTCCACCCGAGCAACAATCTCGAAGATCTCGAA	19455
Qy	561	ValIleThrSerTrpPheGlnSerThrArgSerLysGluAspAlaIleProGluAlaLeu	580
Dp	1946	GTTTACTCTTGCTGGCTTTCAGAGCAGATGAGCAAGAGAGAGCCGATCCGGAAGCACTG	20055
Qy	581	LysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAsnPheLys	600
Dp	2006	AAAAGTCTCATATTCGGAATTTTGAACCTTTGGCAAAATTTCTATCTAATTTTCTCAAG	20655
Qy	601	GluIleGluGlnArgLeuAlaLeuTrpGluArgSerAsnAlaGlnIleArgAspTYr	620
Dp	2066	GAATTTGAGCAACGCACTTGCCCTTGCGGAAGCGCGCTCAATGCCCCAATCAGAGATTAC	21255
Qy	621	GlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysValIleAlaHis	640
Dp	2126	CAAGAAGATCGCGAGATGTATGTCTGAAGAACTTCAGGGCAATGAAGACCTGGCGCTCAC	21855
Qy	641	LeuTrpLysHisSerGluAlaLeuGluValaLeuGluAsnGlyIleLysSerSerArgArg	660
Dp	2186	CTGTGGAAAGCACAGGAGGCTTTGGAGGCGCTTGAGAAATGGAATCGAAGCTCCCGCGCG	22455
Qy	661	LeuGluAsnPheCysArgAspPheGluLeuGlnLysValCysTYrLeuProLeuAsnThr	680
Dp	2246	CTGAGAGAACTTCTGACAGAGACTTTTGAGCTGCAAGAGGTGTTCATCCACGCTCAACACC	23055
Qy	681	PheLeuLeuAspProLeuHisArgLeuMetHisArgTYrLysGlnValLeuGluArgLeuCys	700
Dp	2306	TTCTCTCTCGCGGACCTGCACCGGCTGTACGTGACATCAAGAGAGGTCTGGACGGCTGTGCG	23655
Qy	701	LysHisHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIle	720
Dp	2366	AAACACCAACCCCGGAGCCAGCCGCACTTCACAGGACTCCGAGCGGCTTTGGCAGAGATC	24255
Qy	721	ThrGluMetValaLaglnLeuHisGlyThyMetIleLysMetGluAsnPheGluLysLeu	740
Dp	2426	ACGAGAGATGGTGCACAGACTCCACGGTACGATGATCAAGAGATGAGAAATTTCCAGAAAGCTG	24855
Qy	741	HisGluLeuLysValAspLeuIleGlyIleAspAsnLeuValProGlyArgGluPhe	760
Dp	2486	CACCAAACTCAAGAAAGATTGATTTGGCATTGACATCTTTGGTGTCCGGGAAGGAGATTC	25455
Qy	761	IleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArgMetPhePhe	780
Dp	2546	ATCCGTCTGGGCAAGCTTCAGAAAGCTCTCGGGGAAAGGGGCTTCAGACGCGATGTTCTTC	26055
Qy	781	LeuPheAsnAspValLeuLeuTYrThrSerArgIleuThrAlaSerAsnGlnPheLys	800
Dp	2606	CTGTTTCAACGACGTCTGTCTATACACGAGCGGGGGCTGACGGCTCCAACTCATTTTAA	26655
Qy	801	ValHisGlyGlnLeuProLeuTYrGlyMetThrIleGluGluSerGluAspGluTrpGly	820
Dp	2666	GTCACACGGGCGAGCTCCGCTCTATGCGCATGACATTGAGGAGACCGAAGACAGATGGGGG	27255
Qy	821	ValProHisCysLeuThrLeuArgGlyValArgGlnSerIleIleValAlaSerSer	840
Dp	2726	GTCGCCCACTGCTGACCTTCGGGGGCCAGGGCGAGTCAATCATGTGGCCGCAATTTCT	27855
Qy	841	ArgSerGluMetGluLysTrpValaGluAspIleGlnMetAlaIleAspLeuAlaGluLys	860
Dp	2786	CGGTTCGAGATGAGAAAGTGGGTTGAGGACATCCAGATGGCATTTGACTGGCGGAGAG	28455
Qy	861	SerSerSerProAlaProGluPheLeuAlaSerSerProAspAsnLysSerProAsp	880
Dp	2846	AGCAGACAGCCCCCGCTTGAGTCTTGGCCAGACGCCCCCTGCAACACATCTCCCTGAT	29055

QY	881	GIU1a1aThra1a1aAaPergInGInuSerGIuAaPleuSer1a1aSerAaThSer1eu	900
Db	2906	GAAGCCACCGGGCTGACACGAGAGTACAGAGATACCTCGACCGCTCGCACATCGCTG	2965
QY	901	GIUa1rgInu1a1aProH1saRG1yAaSnThMeTValH1a1aValCybTPH1saRGaSnThr	920
Db	2966	GAGGCCACGGCCCCCGCACCGCGGCACACATGGTGCACGCTGTCTGGCACCGCAACC	3025
QY	921	SeVal1aSerMeTVal1aAaPheSer11e1a1aValGIuAaGIn1eAaSerGIyAaSn1eu	940
Db	3026	AGCGTCTCCATGAGTGAGACTTCAGATCGCAGATGAGATCAGTTGTCTGGAAACCTGCTG	3085
QY	941	Arg1yAaPhe1yAaAaSerAaSnGIyTTrGIn1y1eUrTpVal1a1aPheThAaPheCy	960
Db	3086	AGGAATTCAAAAACAGCAACGGGATGGCAAGCTGGGGTGCTGCACAACTTCCTGC	3145
QY	961	LeuAaPhePheTyr1yAaSerH1eGI1aAaPaaSnH1aPro1eAa1aSer1euPro1eAaGIy	980
Db	3146	CTGTCTCTTCACAAATCACACACAGACATATCCCTTGGCACCTGCCCTCTGCTGGC	3205
QY	981	Ty1Ser1euThr11eProSerGIuAaSn11eGI1yAaP1yTVal1aPhe1yAa1eu	1000
Db	3206	TACTCGCTCACACATCCCTCTGAGTCCGAGAACATCCAGAAAGACTACG1eTTCAAGCTG	3265
QY	1001	H1aPhe1yAaSerH1a1yTyrTyrPha1a1aGIuAaSerGIyTTrH1aPheGIuAaTTrp	1020
Db	3266	CAC TTCACAGTCCACAGCTTACTTACTTACGGCGGAAAGCGAGTACAGTTGAAAGGTGG	3325
QY	1021	MeGIuAa111eAaRGSer1aThSerSer1aSerAaRGProH1a1a1eAaSerH1e1y	1040
Db	3326	ATGGAAGAGATCCCAAGGCCACACAGCTGTCTCGGAGCCCAACGTTGAGGCACAAA	3385
QY	1041	GIuSer1euAa1yTyr 1045	
Db	3386	GAGTCTCTTGTGTAT 3400	
RESULT 4			
LOCUS	CS130723	5028 bp	DNA linear PAT 02-AUG-2005
DEFINITION	Sequence 9 from Patent WO2005064009.		
ACCESSION	CS130723		
VERSION	CS130723.1	GI:7192793	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1	Obernoff, T.F.	
AUTHORS		Classification of cancer	
TITLE		Patent: WO 2005064009-A 9 14-JUN-2005;	
JOURNAL		Aeros Applied Biotechnology APS (DK)	
FEATURES		Location/Qualifiers	
source		1..5028	
ORIGIN		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
		/note="FBRM, RhogEP (ARRGEF) and plectesterin domain pro"	
Alignment Scores:			
Pred. No.:	0	Length:	5028
Score:	5463.00	Matches:	1045
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0
US-09-555-342B-2 (1-1045) x CS130723 (1-5028)			
QY	1	MeGIuAa111eGIuAaRGProThProGIySerAaRG1eGIuAa1aProGIuAaSnSer	20

Db 337 ATGGAGAAATATAGCAGAGCGACCCGACGATCAAGTGGGGGGCCCGAAAAATTGG 396
 Qy 21 GYTLISserThreugluuarglygluylvsProProBorthProserglylylsleuVal 40
 Db 397 GGGATCAGTACCTTGGAACCTGACAGAAACCGCCCCCAACCTTCAGGAAAACTCGTG 456
 Qy 41 SerILysylIleGlnMetIleuAspAspThrGlnGluIlePheGlnValProGlnIAspAla 60
 Db 457 TCCATCAAAATCCAGATGCTGGATGACACCCAGAGGCACTTTGAAGTTCCACAAAGAGCT 516
 Qy 61 ProGlyLyysValIleuIleuAspAlaValCysAsnHisIleuAsnIleuValGluIAspTYr 80
 Db 517 CTGGGAGAGGTGCTGCTGGATGCGAGTTTGCACCACTCAACCTCGGAGAGGTGACTAT 576
 Qy 81 PheGlyIleuGluIleuProAspHisIlylsylValIleThyValIlePheIleuIleuIlePro 100
 Db 577 TTTGGCCTCAGAGTTTCTGTATCAAAAAGATCAGCGGTGGCTGGATCTCTTAAAAACC 636
 Qy 101 IleValIysGlnIleArgAspProLyHisValValIAspPheValValIysPhePhe 120
 Db 637 ATTGTGAACAGATTAGAGGCCAAAGCAGCTGTGTGTTAAGTTTGGTGAATTTCTTT 696
 Qy 121 ProProAspHisThrGlnIleuGlnGluIleuThyArgTYrIleuPheAlaIleuGlnVal 140
 Db 697 CCGCTGACCAACACAACTCCAAGAAAGACTCAAGAGTACCTGTCGCGCTGCAAGTG 756
 Qy 141 LysGlnIleAspIleuIleGlnIleuArgIleuThyCysAsnAspThrSerAlaIleuIle 160
 Db 757 AACCAAGACTTGGCTCAAGGAGAGTGAAGTGAATGACACAGCGAGCTCTTGATT 816
 Qy 161 SerHisIleValGlnIleuGlnIleGlyAspPheAspGluIleuIleuAspArgGlnHisIleu 180
 Db 817 TCACACTTGTGAACTGTGAGATTTGGGGAATTTGATGAAAGCTTTGACAGAGACACTTA 876
 Qy 181 AlaIleAsnIleuTYrIleProGlnIleAspAlaIleuGluIleuIleValGluPheHis 200
 Db 877 GCAAAAATTAATATACATACCTCAGCAAGACGCACTAGAGAGCAAAAATCGGGAATTTAC 936
 Qy 201 HisAsnHisIleIleGlnIleuThyProAlaGlnIleAspPheGlnIleuIleuGlnIleAlaArg 220
 Db 937 CATTAACCACTTGGACAAACACAGCAGAGATTCAGATTTCCAGCTCTAGAGATTCGCCGT 996
 Qy 221 ArgIleuGlnMetTYrGlyIleArgIleuHisIleProAlaIleAspArgGlnIleThyIle 240
 Db 997 CGGCTAGAGATGATGGAAATCCGGTTGCACCCGGCCAAAGACAGGAGGCAAGAGATC 1056
 Qy 241 AsnIleuAlaValIleAsnThrGlyIleIleuValPheGlnIlePheThrIlyIleAsnAla 260
 Db 1057 AATCTGGCCGTTGCCAAACAGGGAATCTAGTGTTCAGGGGTTTCACTAAGATCAATGCC 1116
 Qy 261 PheAsnThrAlaIleuValArgIlyIleuSerPheIleArgIlyIleuIleuIleValIleu 280
 Db 1117 TTCAACTGGGGCAAGGCGGAAAGCTGAGCTTCAAGGAGAAAGCGTTTTCATCAAGCTC 1176
 Qy 281 ArgProAspAlaAsnSerAlaTYrGlnAspThrIleuGlnIlePheIleuIleAlaSerArgAsp 300
 Db 1177 CGGCCAGATGCCAATAGTGGTACAGAGATCCTTGAAATTCCTGATGGCCAGATCCGGAT 1236
 Qy 301 PheCysIleSerPheThrIlyIleCysValGlnHisIleAlaPhePheArgIleuPheGlu 320
 Db 1237 TTCTGCAGAGCTTTCGAAAAATCTGTGTGAACATCATGCTCTTTAAGCTTTTGA 1296
 Qy 321 GluProIleuProIleuProIleuProValIleuPheSerArgIlyIleSerPheAspPheSer 340
 Db 1297 GAACCCCAAAACCAAGCCCAAGCCGTCCTTTAAGCCGGGGGTATCATTTGCTTCAAGT 1356
 Qy 341 GlyArgThrGlnIleuGlnValIleuAspTYrValIleGlnIlyIleIleValIleGln 360
 Db 1357 GGTCCGAGCTCAGAAAGAGGTTCCTGACTATGTTAAAGAGAGGACATAAGAGGTGCGAG 1416
 Qy 361 PheGlnArgIlyHisIleSerIlyIleHisIleSerIleArgSerIleuAlaSerGlnProThrGln 380

Db 1417 TTGAAAGGAGACACAGCAAGATTCAATCTATCCGAGCCTTGCTTCAAGCTTACAGAA 1476
 Qy 381 IleuAsnSerGlnValIleuGlnIleuSerGlnIleuSerThrSerIleuThrPheGlnIlyGlu 400
 Db 1477 CTGAATTCGAAAGTGTGGAGGAGTCTCAGCAGAGCACAGACCTTACATTTGGAGAAAGT 1536
 Qy 401 AlaGlnSerProGlyIleGlnIleuSerCysArgArgGlyIlyGlnIleProIleuValSerAlaGly 420
 Db 1537 GCGGAATCTCCAGGGGGCCAGAGCTCCGGCCGAGGAAAGAAACCGAAAGGTTTCCGCCGG 1596
 Qy 421 GluProGlySerHisIleProSerProAlaProArgIlySerProAlaIleIleuIleGlnAla 440
 Db 1597 GAGCCGGGTGCCACCGGAGCCTCGCGAGAGAAAGCCCGCGGTAAACAAGAGGGG 1656
 Qy 441 AspGlyAlaIleSerAlaProThrGlnIleuGlnIleuValIleAspAspArgThrGln 460
 Db 1657 GACGAGCCGCTCCGGCCGCCAGAGAAAGAGAGAGTGTGTTAAGATAGAGACCCAG 1716
 Qy 461 GlnIleuSerProGlnIleuProGlnIleuProSerThrGlySerIleuThrGlySerProHisIleu 480
 Db 1717 CAGAGTAACTCAGCCGCCGAGCCAGCAACAGAGCTCTGACTGCGAGTCCCTCACCTT 1776
 Qy 481 SerGluIleuSerValIleuSerGlnIlyValAlaProAlaAsnValIleIleuSerPro 500
 Db 1777 TCGAGCTGTCTGTAACTCGAGAGGGGAGTGGCCCTCGCAACCTGACCTTGTCTCC 1836
 Qy 501 AsnIleuSerProAspThrIleuGlnIleuSerProIleuIleSerProIleuIleuAspGln 520
 Db 1837 AACCTAGCCGCCAGCACCAAGAGGCTCTCTCTTATATAGCCCGCTGTGAATGACAG 1896
 Qy 521 AlaCysProArgThrAspAspGluIleuArgGlyArgArgIlyAspPheProThrAspIly 540
 Db 1897 GCTGCCCCCGGAGCAGATGAGATGAGGGCCGGAGGAAAGATTTCCCACTGATAA 1956
 Qy 541 AlaTYrPheIleAlaIleuGlnValSerThrThrGlnIleuArgTYrIleuIlyAspIleuGln 560
 Db 1957 GGTACTCTTATAGCTTAAGAAAGTGTCTTACCAGGAGCAACATCTGAAGATTCGAA 2016
 Qy 561 ValIleThrSerTriPheGlnIleuSerThrValSerIlyGlnAspAlaMetProGluAlaIleu 580
 Db 2017 GTTATACCTTGTGGTGTGATGACAGCAGTGAAGCAAGAGAGCCCATGCCGGAAGACTG 2076
 Qy 581 LysSerIleuIlePheProAspPheGluProIleuHisIlyPheHisIleThrAsnIleuIly 600
 Db 2077 AAAAGTCTATATTCGCCAATTTTGAACCTTGCACAAATTCATCAATTTTCAAG 2136
 Qy 601 GlnIleGlnIleuArgIleuAlaIleuTrpGlnIleuArgSerAsnAlaGlnIleuArgAspTYr 620
 Db 2137 GAAATTGACCAACGACTTCCCTGTGGAAAGCCGCTCAAAATGCCAAATCAGAGATTAC 2196
 Qy 621 GlnArgIleGlyAspValMetIleuIlyAsnIleGlnIlyMetIlyHisIleuAlaHis 640
 Db 2197 CAAGAGATGGGAGATTCATGCTGAAGAACATTCAGGGCATGAAGCACTCGGGGCTCAC 2256
 Qy 641 LeuTrpIlyHisIleSerGlnAlaIleuGlnAlaIleuGlnIlyIleIlySerSerArgArg 660
 Db 2257 CTGTGGAACAACAGCAGAGCCTTGGAGGCCCTGAGAAATGGAATCAAGAGCTCCCGCGG 2316
 Qy 661 LeuGlnAsnPheCysArgAspPheGlnIleuGlnIlyValCysTYrIleuProIleuAsnThr 680
 Db 2317 CTGGAAACCTTCTGCAGAGACTTGAAGTGCAGAAAGTGTGTTACCTTACCGCTCAACACC 2376
 Qy 681 PheIleuIleuArgProIleuHisArgIleuMetHisTYrIlyGlnIleuGlnIleuArgIleu 700
 Db 2377 TTCTCTCTCGGCGACTGACCGGCTCATGCACTACAGACAGTCTTGAGAGCGGCTGTGC 2436
 Qy 701 LysHisIleProProSerHisAlaAspPheArgAspCysArgAlaAlaIleuAlaGluIle 720
 Db 2437 AAACACCAACCCGCGAGCAGCCGACTTCAGAGGATGCGCGAGCCGCTTTGGCAGAGATC 2496
 Qy 721 ThrGlnMetValAlaGlnIleuHisIleGlyThrMetIleuMetGlnIleuAsnIleuIlyIleu 740
 Db 2497 ACGAGAGTGTGCACAGCTCCAGCGGTACGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2556

QY 741 HIGGIGLUTLEUYSYLAAPLEUIGLYILEHAPSNLEUVAL1PROG1YARGLUPHE 760
DB 2557 CAGCACTCAGANAGATTGATTGCAATTCCTGTGGTCCGGGAAGGAGTTTC 2616
QY 761 IIAARGLEUGLYSERLEUSERYLEUSERYLYSGLYLEUGINGINATGMPHEPHE 780
DB 2617 ATCCGCTGGGGCAGCTCAGCAAGCTCCGGGGAGAGGGGCTCCAGCAGCGCATGTTCTTC 2676
QY 781 LEUPHEANAPVALLEULEUYTTHSERARGLYLEUTHRALASERANGINPHALYS 800
DB 2677 CTGTTCACACGCTGCTATACACAGCCGGGGGCTGACGGCTCTCAATCAGTTTAA 2736
QY 801 VALHIGLYGILEUPOLEUYTYRGLMEETHRIILEGUGLUSERYLUAAPGLUTRGLY 820
DB 2737 GTCCACGGGAGCTCCCGCTTATGCAATGACATGAGAGACGAGACGAGTGGGG 2796
QY 821 VALPROHISCYLEUTHRLLEUARGLYGLNARGLINSERILEVAL1A1ASER 840
DB 2797 GTGCCCACTGCTGACCTCCGGGGCAGGGGCAATCATCTGCGCCGCAAGTTCT 2856
QY 841 ARSERYGLMEUGLYVTRTPVALGLUAPRILEGIMETALAL1EAPLEUALAGLULYS 860
DB 2857 CGGTCCAGATGGAGATGGGTGAGACATCCAGATGGCCATTCCTGGCGAGAG 2916
QY 861 SERSERPROVALAPROG1UPHELEUALASERPROVALAPVALYSERPROASP 880
DB 2917 AGCAACAGCCCCCGCTGAGTTCTTGACGACGCCCCCTGACACAGTCTTCGAT 2976
QY 881 GLUALATHRAL1A1AAPGINGLUSERYGLUAPRLEUSERYLASEATGTHSERLEU 900
DB 2977 GAGGCCACCGCGGTGACAGAGATGACAGAGATGACCTGAGCGCTCGCCGACATGCTG 3036
QY 901 GLUARGGLNALAPROHISARGLYAETHMETVALHISVALCYETTPHISARGAETHR 920
DB 3037 GACGCGCAGGCGCCGCGCACCGCGCACACATGTGTGACGTGTCTGCGCACCGCACACC 3096
QY 921 SERVALSERMETVALAPRHSERILEALVAL1G1UANGINLEUSERYLYAETHLEU 940
DB 3097 AGGCTTCCATGTTGATGATTCAGCATGCAAGTGAAGATGATGTTCTGGAACCTTCG 3156
QY 941 ARGLYAPHELYSANSERANGLYTRPG1NLVLEUTHRYVAL1PETHRASNPHCY 960
DB 3157 AGGAATTCACAAAACAGCAACGGGTGACAGAGCTGGGGTTCACAACTTCTGC 3216
QY 961 LEUPHEPHERYLYSERHISGLNAPRASNHISPROLEUALASERLEUPROLEULEGLY 980
DB 3217 CTGTTCCTTACAAATCACACAGAGCAATCATCCCTTGCAGCTGCTGCTGCGC 3276
QY 981 TYRSELEUTHRIEPROSERGLUSERYGLUASNL1EGLINYSAPRYVAL1PHELYLEU 1000
DB 3277 TACTCGCTCACATCCCTCGAGTCCGAGAACATCCAGAAAGCTACGTTTCAACCTG 3336
QY 1001 HISPELYSERHISVAL1TYRTPHEARG1AGLUSERYLYTRHPHEGLUARGTRP 1020
DB 3337 CACTTCAAGTCCACGCTACTACTTCAAGGCGGAAAGCAGATACAGTTGAAAGTTGG 3396
QY 1021 METGLUVAL1LEARGSERALATHRSERSEVALASERAPROHISVAL1LEUSERHISLYS 1040
DB 3397 ATGGAAGATGATCCGAGTGCACACAGCTGCTGCGGACCCGCTGTGATGATCAGAAA 3456
QY 1041 GLUSERYLEUVAL1TYR 1045
DB 3457 GAGTCTCTTGTGTAT 3471

RESULT 5
BD31200 4687 bp DNA linear PAT 17-JUL-2003
LOCUS Human cytoskeleton associated proteins.
DEFINITION BD31200
ACCESSION BD31200.1 GI:33040970
VERSION JP 2002526076-A/2.
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
AUTHORS Lal, P., Tang, T. Y., Yue, H., Hillman, J. L., Bandman, O., Corley, N. C.,
Guegler, K. J., Patterson, C., Azimzai, Y. and Baughn, M. R.
TITLE Human cytoskeleton associated proteins
JOURNAL Patent: JP 2002526076-A 20-AUG-2002;
INCYTE PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002526076-A/2
PD 20-AUG-2002
PR 17-SEP-1999 JP 2000574254
PR 18-SEP-1998 US 60/172226-27-APR-1999 US 60/131321 PI
PREPRINT LAL, TOM Y TANG, HENRY YUE, JENNIFER L HILLMAN, OLGA PI
BANDMAN,
PI NEIL C CORLEY, KARL J GUEGLER, CHANDRA PATTERSON, YALDA AZIMZAI,
PI MARIAN R BAUGHN
PC C12N15/09, A61K38/00, A61K45/00, A61P1/16, A61P3/00, A61P3/06 PC
A61P3/10, A61P5/00,
PC A61P5/38, A61P7/00, A61P7/06, A61P9/00, A61P9/10, A61P11/06 PC
A61P13/00, A61P13/12,
PC A61P17/06, A61P19/02, A61P19/10, A61P21/04, A61P25/00, A61P25/00,
PC A61P25/04,
PC A61P25/14, A61P25/16, A61P25/28, A61P29/00, A61P31/04, A61P31/12,
PC A61P33/00, A61P35/00, A61P35/02, A61P37/00, A61P43/00, C07K14/47,
PC C07K16/18,
PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12P21/08 PC
, C1201/68, G01N33/15,
PC G01N33/50, G01N33/53, G01N33/53, G01N33/566, C12N15/00, C12N5/00,
PC A61K37/02
CC Incyte ID No: 1320252
FH Key Location/Qualifiers
FT source 1. 4687
1. 4687
/organism='Homo sapiens (human)'.
FEATURES
source location/Qualifiers
1. 4687
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 4687
Score: 5459.00 Matches: 1044
Percent Similarity: 99.9% Conservative: 0
Best Local Similarity: 99.9% Mismatches: 1
Query Match: 99.9% Indels: 0
DB: 2 Gaps: 0
US-09-555-342B-2 (1-1045) x BD31200 (1-4687)
QY 1 METGLUVAL1LEUARGSERALATHRSERSEVALASERAPROHISVAL1LEUSERHISLYS 20
DB 90 ATGGAGCAATTAAGCAGAGGCGGACCCCGAGATCAGACCTGGGGGCGCGGAAAAATTGG 149
QY 21 GLYIIESERTHRIEUGLARGGLYGLNLYSAPROPROBOTHTRPROSERGLYVLEUVAL 40
DB 150 GGGATCAGTACTTGGAAACGTGACAGAAACCGCCCAACACCTTCAAGAAAACTGCTG 209
QY 41 SERILEUYRILEGIMETLEUAPRTHRG1GLUVAL1HAGL1VAL1PROG1NARGVAL 60
DB 210 TCCATCAAAATCCAGATGCTGATGACACCCAGAGGCACTTTAAGTTCCACAAAGAGCT 269
QY 61 PROG1LYSVAL1LEUAPRTHRG1GLUVAL1HAGL1VAL1HAGL1VAL1HAGL1VAL1HAGL1 80
DB 270 CTGGGAGAGGTGCTGCTGATGACAGTTGCAACCACTCACTCGTGAAGGATGACTAT 329
QY 81 PHEGLYLEUGLUPHEPROAPRHSISLYSGLYRILETHRYVAL1TRPLEUAPLEULEUYSPRO 100
DB 330 TTGGGCTCGAGTTTCTGATCACAAAAGATCAGGATGTGGCTGATCTCTTAAAAACC 389

QY 101 ILeValIysGlnIleArGArProLySHLeValIValIySPheValIlySPhePhe 120
DB 390 ATGTGTAAACAGATTAGAGGCCAAAGCACGTTGTCTTTAAAGTTTGTGGTAAATCTTT 449
QY 121 ProProAspHLeThrgInLeuGInGluIleuThArGTYrLeuPheAlaLeuGInVal 140
DB 450 CCGCTGTACACACAACTCCAAAGAAAGACTCAAGGTACTGTTCCGGCTGCAGGTG 509
QY 141 IyGGLnAspLeuAlaGInGlyArGLeuThCyAspAspThrSerAlaAlaLeuLeuIle 160
DB 510 AACAGAGACTGTGGCTCAAGGAGAGTGAAGTGAATGACACAGCGAGCTCTTGATT 569
QY 161 SerHLeIleValGInSerGluIleGlyAspPheAspGluAlaLeuAspArGInHLeu 180
DB 570 TCACACTTGTGCATCTGAGATTGGGGAATTTGATGAACCTTGGACAGAGACTTA 629
QY 181 AlaLyAspLeuIleTYrIleProGInGInAspAlaLeuGInAspLyIleValGluPheHLe 200
DB 630 GCAAAATATTAATACATACCTCAGCAAGACGCACTAGAGCAAAATCGTGAATTTTCA 689
QY 201 HLeAspHLeIleGlyGInThrProAlaGInSerAspPheGInLeuGInIleAlaArG 220
DB 690 CATAGCACATTGACAAACACACAGAGATCAAGATTTCCAGCTCTAGAGATTGCCGT 749
QY 221 ArGLeuGInMetTYrGlyIleArGLeuHLeProAlaLyAspArGInGlyThrLyIle 240
DB 750 CCGCTAGAGATGTATGGAAATCCGGTTCACCCGGCCAAAGACAGGAGGACAGAGATC 809
QY 241 AspLeuAlaValAlaAspThrGlyIleLeuValPheGInGlyPheThrLyIleAspAla 260
DB 810 AATCTGGCCGTTGCCAAACAGGGAATTTCTAGTCTTTCAGGGTTTCACTAAGATCAATGCC 869
QY 261 PheAspThrAlaLyAspValaArGlyLeuSerPheLyAspArGlyPheLeuIleLyLeu 280
DB 870 TTCAACTGGGCGCAAGGCGGAAGCTGAGCTTCAAGAGGAAGCGCTTTTCAACAAGCTC 929
QY 281 ArGProAspAlaAspSerAlaTYrGInAspThrLeuGInPheLeuMetAlaSerArGAsp 300
DB 930 CCGCCAGATGCCAATAGTGGCTACAGAGATCCTTGGAAATTCCTGATGGCCACTCCGGAT 969
QY 301 PheCyAlaSerPheThrLyIleCyValaGluHLeHLeAspPheArGLeuPheGlu 320
DB 990 TTCTGAAAGCTTCTGGAAATCTGTGTGAACATCATGCTCTTTAGACTTTTGGAA 1049
QY 321 GluProLyProLyProLyAspProValLeuPheSerArGlySerSerPheArGpPheSer 340
DB 1050 GAAGCCCAACCAAGGCCCAAGCCGCTCTCTTAAAGCGGGGTCATCATTTCCGTTCAAGT 1109
QY 341 GlyArGThrgInLyGInValIleuAspTYrValIyGInGlyIyHLeIyLyIyValaGIn 360
DB 1110 GGTGGAGCTCAGAGAGAGGTTCTCGACTATGTTAAAGAGAGAGACTTAAGAGAGTGCAG 1169
QY 361 PheGluArGlyLeHLeSerLyIleHLeSerIleArGSerLeuAlaSerGInProThrglu 380
DB 1170 TTTGAAGAGAGACAGCAAGATTCTATCTATCCGGAGCTTCTCACAGCTTACAGAA 1229
QY 381 LeuAspSerGluValIleuGInGInSerGInGInSerThrSerLeuThrPheGlyGInGly 400
DB 1230 CTGAATTCGAAATGCTGAGAGCTCAGACAGACACAGCTTACATTGAGAGAGGT 1289
QY 401 AlaGluSerProGlyGInSerCyAspArGlyLyGInGluProLyAspValaSerAlaGly 420
DB 1290 GCGGAATCTCCAGGGGGCCAGAGCTGCGGCGAGAGAAAGAACGAAAGTTTCCGCGGG 1349
QY 421 GluProGlySerHLeProSerProAlaProArGArGSerProAlaGlyAspLyGInAla 440
DB 1350 GAGCCCGGGGTGCACCGAGCTTGGCGCCAGAGAGAAAGCCCGCGGTAAACAAGCGGGCG 1409
QY 441 AspGlyAlaAlaSerAlaProThrgInGluGInGluValaValaLyAspArGThrgIn 460
DB 1410 GAGGAGCGCGCTGGGGCCCAAGAGAGAGAGAGGTCTGTAAAGATAGACCCAG 1469

QY 461 GInSerLyProGInProProGInProSerThrgLySerLeuThrgLySerProHLeu 480
DB 1470 CAGAGTAAACTCAGGCCCCGAGAGCCAGACAGAGCTCCCTGACTGGAGTCTCACTT 1529
QY 481 SerGInLeuSerValaAspSerGInGlyValaAlaProAlaAspValaThrLeuSerPro 500
DB 1530 TCCGAGCTGTCTGTAACTGCAGGGGGAGTGGCCCCCTGCACCACTGTGTCTCCC 1589
QY 501 AspLeuSerProAspThrLyGInAlaSerProLeuIleSerProLeuAspAspGIn 520
DB 1590 AACCTAGCCCCGACACCAAGAGAGCTCTCCCTTATCAGCCGCTGTAATGAGCCAG 1649
QY 521 AlaCyProArGThzAspAspGluAspGInGlyArGArLyAspArPheProThzAspLy 540
DB 1650 GCCTGCCCGGAGCGAGCATGAGATGAGGGCCGAGAGAAAGATTCCCAATGATAA 1709
QY 541 AlaTYrPheIleAlaLyGInValaSerThrThrgInArGThrTYrLeuLyAspLeuGlu 560
DB 1710 GGTACTTCAATAGCTAAGAGAGTGTCTACACCGAGCAACATATCTGAAGATCTCGAA 1769
QY 561 ValIleThrSerThrPheGInSerThrValaSerLyGluAspAlaMetProGluAlaLeu 580
DB 1770 GTTATCACTTGTGTGTTTCAAGACAGATGAGCAAGAGAGAGAGCCATGCGGAAAGACTG 1829
QY 581 LySerLeuIlePheProAspPheGluProLeuHLeIyPheHLeThzAspPheLeuLy 600
DB 1830 AAAAGTCTCATATTCGCCAATTTTGAACCTTTGCAAAATTTCAATATTTTCCAG 1889
QY 601 GluIleGInGInArGLeuAlaLeuThrgInGlyArGSerAspAlaGInIleArGAspTYr 620
DB 1890 GAAATTGACCAACGACTTCCCTGTGGAGAGCGCGCTCAAAATGCCAAATTCAGAGATTAC 1949
QY 621 GInArGIlleGlyAspValaMetLeuLyAspAlaGInGlyMetLyHLeIleuAlaHLe 640
DB 1950 CAAGATTCGGGATTCATGCTGAAGACATTCAGGGCATGAAGACCTGGCGGCTCAC 2009
QY 641 LeuThrLyHLeSerGInAlaLeuGInAlaLeuGInAspGlyIleLySerSerArGArG 660
DB 2010 CTGTGAACACACAGAGGCTTTGAGGCTCGTGAAGATGGAATCAAGAGCTCCGGCGG 2069
QY 661 LeuGInAspPheCyArGAspPheGInLeuGInLyValaCyTYrLeuProLeuAspThr 680
DB 2070 CTGGAGAACTTCTGCAGAACTTTGAGCTGCAAGAGTGTGTTACTTACCGCTCAACACC 2129
QY 681 PheLeuLeuArGProLeuHLeArGLeuMetHLeTYrLyGInValaLeuGInArGLeuCy 700
DB 2130 TTCTCTCTCGGCGCACTGACCGGCTCATGCACTACAGACAGAGTCTTGAAGCGGCTGTGC 2189
QY 701 LyHLeHLeProProSerHLeAlaAspPheArGAspCyArGAlaAlaLeuAlaGluIle 720
DB 2190 AAACACACACCGCGCAGACCGCACTTCAAGGACTGCCAGCGCTTTGGCAGAGATC 2249
QY 721 ThrgInMetValaAlaGInLeuHLeGlyThrMetIleLyMetGInAspPheGInLyLeu 740
DB 2250 ACGAGATGTGGCACAGCTTCAAGGTAGATCAAGATGAGAAATTTCCAGAAAGCTG 2309
QY 741 HLeGInLeuLyAspLeuIleGlyIleAspAspLeuValaProGlyArGInGluPhe 760
DB 2310 CAGAACTCAAGAAAGATTGATTGGCATTTGACATCTGTGTCTCCGGAAGAGAGTTC 2369
QY 761 IleArGLeuGlySerLeuSerLyLeuSerGlyLyGlyLeuGInGInArGMetPhePhe 780
DB 2370 ATCCGTCTGGGACGCTCAGCAAGCTCTCGGGGAAGGGGCTCCAGACGCGCATGTTCTTC 2429
QY 781 LeuPheAspAspValaLeuLeuTYrThrSerArGlyLeuThrAlaSerExtGInPheLy 800
DB 2430 CTGTTCAAGAGAGTCTGTATTAACAGAGCGGGGCTTACAGGCTTCAATCAGTTTAA 2489
QY 801 ValHLeGlyGInLeuProLeuTYrGlyMetThrIleGInGlySerGInuAspGluTYrGly 820
DB 2490 GTCCAGGGGAGACTCCGCTCATGTGCATGACATTTAGAGAGAGCAACAGAGTGGGG 2549
QY 821 ValProHLeCyAspLeuThrLeuArGlyGInArGInSerIleIleValaAlaAspSer 840


```

Db      2550 GTGCCCCACCTGCTGACCCCTCCGGGCGCAGCGCATCATCTGCGCCGCACTTCT 2609
Qy      841 ArgSerGluMetGluValTyrPValGluAapIleGlnMetAlaIleAapLeuAlaGluLys 860
Db      2610 CGGTCCCAAGATGGAAGATGGGTGTGAGGACATCCAGATGGCCATTGACCTGCGCGAGAG 2669
Qy      861 SerSerSerProAlaProGluLysPheLeuAlaSerSerProProAlaAapLysSerProAap 880
Db      2670 AGCAGCAGCCCCGCGCTGAGATCTCTGGCGCAGAGCCCCCTGACAAACATCTCCCTGAT 2729
Qy      881 GluAlaThrAlaAlaAapGlnGlnLysSerGluAapPheLysSerAlaSerArgThrSerLeu 900
Db      2730 GAAGCCACCGCGGCTGACAGAGATCAAGAGATACCTGACCGCTCGCGACATCTGCTG 2789
Qy      901 GluArgGlnAlaProAlaArgGlyAapThrMetValHisValCysThrPheAapAapThr 920
Db      2790 GAGCGCAGAGGCGCGCAGCCGCGCAACATATGGTGACGTTGCTGGCACCGCAACACC 2849
Qy      921 SerValSerMetValAapPheSerIleAlaValGluAapGlnLysSerGlyAapLeuLeu 940
Db      2850 AGCGTCTCCATGATGAGACTTACGATCGCATGGAGATCATGTTGTCTGGAACCTGCTG 2909
Qy      941 ArgLysPheLysAapSerAapGlnLysTyrGlnLysLeuTyrValPheThrAapPheCys 960
Db      2910 AGGAAATTCACAAAACAGCAACGGGTGCGAGAGCTGTGGGTGTCTTACAAACTTCTGC 2969
Qy      961 LeuPhePheTyrLysSerHisGlnAapAapHisAproLeuAlaSerLeuProLeuLeuGly 980
Db      2970 CTGTCTTCTTCAATATCACACACAGACATATCATCCCTTGGCCAGCTGCTGCTGCGG 3029
Qy      981 TyrSerLeuThrIleProSerGlnSerGlnLysGlnLysAapTyrValPheLysLeu 1000
Db      3030 TACTCGCTCACCACTCCCTCTGAGTCCGAGAACTCCAGAAAGACTAGTGTCAACTG 3089
Qy      1001 HisPheLysSerHisValTyrTyrPheArgAlaGlnSerGluTyrTyrPheGluArgTyr 1020
Db      3090 CACTTCAAGTCCCACTGCTACTTACTTCAAGGCGGAAAGCGAGTACAGCTTGAAGGTGG 3149
Qy      1021 MetGluValIleArgSerAlaThrSerSerAlaSerArgProHisValLeuSerHisLys 1040
Db      3150 ATGGAATGTATCCGACAGTGCACACAGCTCTGCTGCGACCCAGTGTGATGATCACAAA 3209
Qy      1041 GluSerLeuValTyr 1045
Db      3210 GAGTCTTGTGTAT 3224

```

```

RESULT 6
BC071592      4932 bp      mRNA      linear      PRI 28-JUN-2005
LOCUS      BC071592
DEFINITION      Homo sapiens FERM, RhogEF (ARHGEF) and plecterin domain protein 1
                (chondrocyte-derived), mRNA (cDNA clone MGC:87400 IMAG:30342239),
                complete cds.
ACCESSION      BC071592
VERSION      BC071592.1 GI:47940451
SOURCE      MGC.
ORGANISM      Homo sapiens (human)
                Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                Homiidae; Homo.
REFERENCE      1 (bases 1 to 4932)
AUTHORS      Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
                Klauener,R.D., Collins,L., Wagner,L., Shenmen,C.M., Schuler,G.D.,
                Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
                Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heideh,F.,
                Diatchenko,L., Marusik,K., Farmer,A.A., Rubin,G.M., Hong,L.,
                Spatton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
                Scheetz,T.E., Brownstein,M.J., Uedin,T.B., Toshiyuki,S.,
                Abramson,J., Piatige,C., Kane,S.S., Loquellano,N.A., Peters,G.J.,
                Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
                McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
                Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

```

```

FEATURES
source
gene
CDS

```

Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
 Scherch,A., Schein,J.E., Jones,S.J. and Marz,M.A.
 Mammalian Gene Collection Program Team
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 2 (bases 1 to 4932)

NIH MGC Project
 Direct Submission
 Submitted (01-JUN-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Stefan Hansson
 CDNA Library Preparation: Michael Brownstein / Ted Usdin
 Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAX Plate: 168 Row: 1 Column: 8
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 48928036.
 Location/Qualifiers
 1..4932
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:87400 IMAG:30342239"
 /issue_type="Placenta, normal"
 /clone_id="NIH_MGC_147"
 /lab_host="DH10B"
 /note="Vector: pBluescriptR"
 1..4932
 /gene="FARP1"
 /note="synonyms: CDEP, MGC87400, PLEKHC2"
 /db_xref="GeneID:10160"
 /db_xref="MIM:602654"
 240..3470
 /gene="FARP1"
 /codon_start=1
 /product="FARP1 protein"
 /protein_id="AAH1592.1"
 /db_xref="GI:47940452"
 /db_xref="GeneID:10160"
 /db_xref="MIM:602654"
 /translation="MGEIQRPTGSRRLGAPENSGISTLERGGKPPPTSPGKLVSTI
 QMLDPOAEVYRQAPRGKTLDAVCNHLNVGQDVGLEFPDHKITWLDLKLIV
 KQIRPRHVAVKVFPPPDITQLEBELRYLPAIOVKDLDLQGRITCDTSAALLI
 SHLVQSEIIGDFDLALREHLAKKTYLPQDALDEKLVFPHNNIGTPASDPQLLEI
 ARLEWYGIKRLADKREGTKIMLVANGILVFOGFTKIMAFNMAKVKRLSPKRFK
 LILKRPDANSAYVDTELEASRDPEKSWKICVEHNAFRLTEEPKPKPVLFRSG
 SSFRSGRTQKQVLDVKEGKKVQPERKSHIHSIRLSAQPTLELNSVLESOOS
 TSLFRPGALSPGSGSCRGKEPKVSGEGRSHSPAPRRSPAGNKOADAAPPE
 EEEVVDRTQKSPQRPQPSSTLSGPHSELSVNSQGVAAANTLSPNLSPPDKO
 ASPLISPLNDQACPTDDEBGRKRPFDKVFYLAKEVSTTERTYLQLEITVSNF
 QSVSVEKDAPPELAKSLIPENPEKLFPHNFKETLEQRLALWEGSNQIROYQRG
 DWMLKTIQCKHILAHIMKYSALALENGIKSRLENPCRPDELQKCYLPLKPTL
 LRPLRLHMYKQYLERLCKHNPSSHADFRCAALAEITEMVAQLHGTMIKMNPFQL

HEBKDLIGIDNLVPEPGPSFSLMRTPLGLQARRIPCAPERRLLLVKEFRLSL
KLSGKLQORMFLLFNDVLLYTSRGLTASQFVHGOLPLYGTTISESEBEMVPL
TLRGOSQIIVASRSSEMEKVEDIOMALDLAKSSSPEDLASSPDNKSPPD
ADQSESDLSASRTSLERQAPHRGNTMTHVCHRRRTSYSMNPSIAVENQSLNR
LKFNSNGWQCLWVFTNPLCLFPYKSHODNHLPLASLGLSLTPESENIQDYVK
LHPKSHVYFRABSEVTFERMMEVIRSATSSASRPHVLSHKESLVV"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	4932
Score:	5431.50	Matches:	1044
Percent Similarity:	97.1%	Conservative:	0
Beet Local Similarity:	97.0%	Mismatches:	0
Query Match:	99.4%	Indels:	31
DB:	5	Gaps:	1

US-09-555-342B-2 (1-1045) x BC071592 (1-4932)

Qy 1 MetGlyGluIleGluGlnArgProThrProGlySerArgLeuGluValaProGluAenSer 20
Db 240 ATGGGGAATAAGACAGAGGCCGACCCAGATCAGACTGGGGGCCCGGAAAATTCG 239
Qy 21 GlyIleSerThrLeuGluArgGlyGlnLysProProThrProSerGlyLysLeuVal 40
Db 300 GGGATCAGTACTTGGAACTGGACAGAAAGCCGCCCAACCTTCAGGAAAACCTCGTG 359
Qy 41 SerIleValIleGlnMetLeuAspAspThrGlnGluValaPheGlnValaProGlnArgAla 60
Db 360 TCCATCAAAATCCAGATGCTGGAATGACACCCAGAGGCAATTTGAAAGTTCCAAAGAGCT 419
Qy 61 ProGlyLysValIleLeuLeuAspAlaValaCysAenHisLeuAenLeuValaGluGlyAspTyr 80
Db 420 CCGGGAGAGTGCTGCTGATGATGAGTTTGCAACCACTCAACCTCGGAAAGGTGACTAT 479
Qy 81 PheGlyLeuGluGlnPheProAspHisLysLysIleThrValaIlePheLeuLeuLysPro 100
Db 480 TTGGGCTCCAGTTCCTGTGATCAAAAGATCAGCGTGTGGCTGATCTCTTAAAAACC 539
Qy 101 IleValIleGlnIleArgAspProLysHisValaValaLysPheValaLysPhePhe 120
Db 540 ATTGTGAACAGATTAGAAAGCCAAAGCAGCTGTGTGTTAAAGTTGTGGTGAATTCCTT 599
Qy 121 ProProAspHisThrGlnLeuGlnGluGlnLeuThrArgTyrLeuPheAlaLeuGlnVal 140
Db 600 CCGCTGACCAACAACAACCTCCAAGAAAGACTCAAGAGTACCTCGGCTGACAGGTG 659
Qy 141 LysGlnAspLeuAlaGlnGlyArgLeuThrCysAenAspHisSerAlaAlaLeuLysIle 160
Db 660 AACCAGGACTTGGCTCAAGGACAGTTGACGTGTAATGACACAGCGAGCTCTCTGATT 719
Qy 161 SerHisIleValaGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgGluHisLeu 180
Db 720 TCCACACTTGTGCAATTCGATTTGGGGAATTTGATGAACCTTGGACAGAGACCTTA 779
Qy 181 AlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIleValaGluPheHis 200
Db 780 GCAAAAAATTAATACATCACTCAAGCAAGCAGCACTAGACGCAAAATCGTGAATTCAC 839
Qy 201 HisAenHisIleGlyGlnThrProAlaGlnSerThrPheGlnLeuLeuGluIleAlaArg 220
Db 840 CAATTAACAACCTTGGACAAACACAGCAGATCAGATTCAGCTCTCTAGAGATTCGCCGT 899
Qy 221 ArgLeuGlnMetTyrGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysIle 240
Db 900 CGGCTAGAGATGATGAAATCCGGTTGACCCGGCCAAAGACAGGAGAGGCAAGAAATC 959
Qy 241 AsnLeuAlaValaAenThrGlyIleLeuValaPheGlnGlyPheThrLysIleAsnAla 260
Db 960 AATCTGGCCGTGGCAACACAGGAAATCTAGTGTTCAGGGTTTCACATAAGATCAATGCC 1019
Qy 261 PheAenThrAlaLysValaArgLysLeuSerPheLysArgLysArgPheLeuIleLysLeu 280
Db 1020 TTCAACTGGGCGCAAGGTGGAAAGCTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTC 1079

Qy 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAlaSerArgAsp 300
Db 1080 CGGCGAGATGCCAATAGTCGTAACAGATACCTTGAATTCCTGATGCCAGCTCGGAT 1139
Qy 301 PheCysLysSerPheThrLysIleCysValaGluHisAlaPhePheArgLeuPheGlu 320
Db 1140 TTCTGCAAGTCTTCTGAAAAATCTGTGTAACATCATGCTTCTTAGACTTTTGGAA 1199
Qy 321 GluProLysProLysProLysProValaLeuPheSerArgGlySerSerPheArgPheSer 340
Db 1200 GAGCCCAAAACCAAGCCCAAGCCGCTCTTAAAGCCGAGGATCATCATTCGTTCCAGT 1259
Qy 341 GlyArgThrGlnLysGlnValaLeuAspTyrValaLysGluGlyLysIleLysLysValaGln 360
Db 1260 GGTCCGACTCAGAAAGAGGTTCTCCGACTATGTTAAAGAAAGAGGACATAAGAGGTGCG 1319
Qy 361 PheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
Db 1320 TTGAAAGAGAGCAGCAAGATTCATTCATCCGAGCCTTGCTCACAGCCTTACAGAA 1379
Qy 381 LeuAenSerGluValaLeuGlnGlnSerGlnGlnSerThrSerLeuThrPheGlyGly 400
Db 1380 CTGAATTCGAAAGTGTGAGCAGTCTCAGCAGAGCACAGCTTACATTTGAGAGAGT 1439
Qy 401 AlaGlnSerProGlyGlyLysSerCysArgArgGlyLysGluProLysValaSerAlaGly 420
Db 1440 GCCGATCTCCAGGGGGCCAGAGCTCCGCGCAGAGAAAGAACCAAGGTTCCGCGCG 1499
Qy 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAla 440
Db 1500 GAGCCGGGGTCCGACCCGAGCCCTGCGCAGAGAGAAAGCCCGCGGTAAACAGAGCGG 1559
Qy 441 AspGlyAlaAlaSerAlaProThrGlnGluGlnGluValaValaLysAspArgThrGln 460
Db 1560 GAGGAGCCGCTCGGCGCCACGAGAAAGAGAGAGAGTCTTAAAGATAGGACCCAG 1619
Qy 461 GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeu 480
Db 1620 CAGAGTAAACCTCAGCCCGCAGCCCAACAGCAGCTCCCTGATCAGGAGTCTCACTT 1679
Qy 481 SerGlnLeuSerValaAsnSerGlnGlyValaAlaProAlaAsnValaThrLeuSerPro 500
Db 1680 TCCGAGCTGTCTGTAACTCGCAGGGGGAGAGTGGCCCTCGCCAACTGATCTGTCTCC 1739
Qy 501 AsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuLeuAsnAspGln 520
Db 1740 AACCTAGCCCGACACCAAGAGGCTCTCCCTTATCAGCCGCTGATGACAG 1799
Qy 521 AlaCysProArgThrAspAspGluAspGluGlyArgArgLysArgPheProThrAspLys 540
Db 1800 GCCTGCCCGGACGAGCATGATGAGTGAAGGGCCGAGAAAGATTCCTCAATGATTA 1859
Qy 541 AlaTyrPheIleAlaLysGluValaSerThrThrGlnArgThrTyrLeuLysAspLeuGlu 560
Db 1860 GCCTACTCATATGCTTAAGAAAGTGTCTACCCAGCAGCAACATATCTGAAGATTCGAA 1919
Qy 561 ValIleThrSerThrPheGlnSerThrValaSerLysGluAspAlaMetProGluAlaLeu 580
Db 1920 GTTATCATCTCGGATTCAGACAGCATGAGCAAGAGGACCCATGCCGAAAGCATCTG 1979
Qy 581 LysSerLeuIlePheProAspPheGluProLeuHisLysPheHisThrAsnPheLeuLys 600
Db 1980 AAAAGTCTATATTCGCAATTTTGAACCTTTGCAAAATTTCTACTATATTTTCCAAG 2039
Qy 601 GluIleGluGlnArgLeuAlaLeuThrGluGlyArgSerAsnAlaGlnIleArgAspTyr 620
Db 2040 GAAATTAGCAACAGATTCCTCGTGGAGAAAGCCGTCAAAATGCCAAATTCAGAGATTAC 2099
Qy 621 GlnArgIleGlyAspValaMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaAlaHis 640
Db 2100 CAAGAAATGGGGATGTCTAGTGAAGAAATTCAGGGCATGAAGCACTGGCGGCTCAC 2159

```

Qy 641 LeuTrpLysHisSerGluAlaLeuGluValLeuGluLysSerSerArgArg 660
Db 2160 CTGGAGAGTACACGAGGCTTGGAGGCTTGGAGAGTACAGAGGCTCCCGCGG 2219
Qy 661 LeuGluAsnPheCysArgAspPheGluLeuGluLysValCysTyrLeuProLeuAsnThr 680
Db 2220 CTGGAGAACTTCTCAGAGACTTTCAGCTGCGAGAGGAGTGTATCTTCAACGCTCAACACC 2279
Qy 681 PheLeuLeuArgProLeuHisArgLeuMetHisTyrTyrGlnValLeuGluArgLeuCys 700
Db 2280 TTCTCTCGGGCGGCACGACCGGCTCATGCAACAGCAGCGCTCGAGCGGCTGTGC 2339
Qy 701 LysHisAspProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIle 720
Db 2340 AACACACACCGCGGAGCCGACCGGACTTCAGGAGCTGCGAGCCGCTTGGCAGAGATC 2399
Qy 721 ThrGluMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPheGlnLysLeu 740
Db 2400 ACGGAGATGGTGGCAGAGCTCCAGGTACGATGATCAAGATGAGAAATTCACAGAGCTG 2459
Qy 741 HisGluLeuLysValAspLeuIleGlyIleAspAsnLeuValValProGlyArg----- 758
Db 2460 CAGCACTCAAGAAAGATTGATGGCAATGCAATCTTGTGTTCCGGGAAGGCTGTGC 2519
Qy 758 ----- 758
Db 2520 TCCTTTCCCTTATGAGAGCTCCACACTTGGAGACAGCAGCGCATCCCATGTGCTCCA 2579
Qy 759 -----GluPheIleArgLeuGlySerLeuSerLysLeu 769
Db 2580 GAGAGAGGCGCATGTTATTAGTAAGAGATTGATCCGTGCGGACAGCTCAGACAGACTC 2639
Qy 770 SerGlyLysGlyLeuGlnGlnArgMetPhePheLeuPheAsnAspValLeuLeuTyrThr 789
Db 2640 TCGGGGAGGGGCTCCAGACGCGCATGTCTTCTGTTCACGACGCTCGCTATACAG 2699
Qy 790 SerArgGlyLeuThrAlaSerAsnGlnPheLysValHisGlyGlnLeuProLeuTyrGly 809
Db 2700 AGCCGGGGGCTGACGGCTTCATCACTTTAAATTCACGGGCGACTCCGCTTATAGGC 2759
Qy 810 MetThrIleGluLysSerGluAspGluTyrPglValProHisCysLeuThrLeuArgGly 829
Db 2760 ATGACGATTAGGAGAGCAGAGAGAGTGGGGGGTGGCCCATCTCGACCTCCGGGGGC 2819
Qy 830 GlnArgGlnSerIleIleValAlaAlaSerSerArgSerGluMetGlyLysTyrValGlu 849
Db 2820 CAGCGGCGATCCATCATCTGTGGCCGCACTTCGCTCCGAGTGAAGAGTGGGTGAG 2879
Qy 850 AspIleGlnMetAlaIleAspLeuAlaGluLysSerSerProAlaProGluPheLeu 869
Db 2880 GACATCCAGATGGCATTTGACCTTGGCGAGAGAGCAGACCCCGCCCTGAGCTTCCTG 2939
Qy 870 AlaSerSerProAspAsnLysSerProAspGluAlaThrAlaAlaAspGlnLysSer 889
Db 2940 GCCAGCAGCCCGCTGACCAAGACTCCCTGATGAAGCAGCCGGGTGACAGAGATCA 2999
Qy 890 GluAspAspLeuSerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgGlyAsn 909
Db 3000 GAGGATACCTGAGCGCTCGCGCACATCGCTGAGGCGCCGCGCCGACCCGGGAGAC 3059
Qy 910 ThrMetValHisValCysTrpHisArgAsnThrSerValSerMetValAspPheSerIle 929
Db 3060 ACAATGGTGCACGTGTCTGGACCCGACACCGAGCTTCATAGTGAATTCAGCATC 3119
Qy 930 AlaValGluAsnGlnLeuSerGlyAsnLeuLeuArgLysPheLysAsnSerAsnGlyTyr 949
Db 3120 GCACTGAGATCAGTGTCTGGAACCTGCTGAGGAATTCAAAACAGCAAGCGGTGG 3179
Qy 950 GlnLysLeuTrpValValPheThrAsnPheCysLeuPhePheTyrLysSerHisGlnAsp 969
Db 3180 CAGAAAGTGTGGGTGTTCACAAACTTGTGCTGTCTTCAAAATCAACACAGAGAC 3239
Qy 970 AsnHisProLeuAlaSerLeuProLeuGluLysSerLeuThrIleProSerGluSer 989

```

```

Db 3240 AATATCCCTTTCGACGCTGCTCTGCTGCTACGCTACATCCCTCGAATCC 3299
Qy 990 GluAsnIleGlnLysAspTyrValPheLysLeuHisPheLysSerHisValTyrTyrPhe 1009
Db 3300 GAGAACATCCAGAAAGACTACGTGTTCAAGCTCAGTCAAGTCCACGCTACTACTTC 3359
Qy 1010 ArgAlaGluSerGluTyrTrpPheGluArgTrpMetGluValIleAspSerAlaThrSer 1029
Db 3360 AGGCGGAAAGCAGATGACGTTGAAAGGTGATGAGATGATCCGACGTGCCACAGC 3419
Qy 1030 SerAlaSerArgProHisValLeuSerHisLysGluSerLeuValTyr 1045
Db 3420 TCTGCTCGGAGACCCACGTTGATGACAAAGAGTCTTGTGAT 3467

RESULT 7
LOCUS CO724059 2937 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 9993 from Patent WO02068579.
ACCESSION CO724059
VERSION CO724059.1 GI:42284916
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 9993 06-SEP-2002.
FEATURES
source 1..2937
location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2937
Score: 4095.50 Matches: 825
Percent Similarity: 79.3% Conservative: 4
Best Local Similarity: 78.9% Mismatches: 32
Query Match: 75.0% Indels: 185
DB: Gaps: 7
US-09-555-342B-2 (1-1045) x CO724059 (1-2937)
Qy 58 GlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGlu 77
Db 1 CAAGAGACTCTGGGAAAGGTGCTGATGATGCTTGGCAACCACTCACTGCTGAA 60
Qy 78 GlyAspTyrPheGlyLeuGluLysPheProAspHisLysValIleThrValTrpLeuAspLeu 97
Db 61 GGTGACTATTGTGACCTCGAGTTCTGATCADAAGAAAGATCAG----- 105
Qy 98 LeuLysProIleValLysGlnIleArgArgProLysHisValValLysPheValVal 117
Db 106 -----GCCAGCTTCCCGGTG 120
Qy 118 LysPhePheProProAspHisThrGlnLeuGlnGluLeuThrArgTyrLeuPheAla 137
Db 121 ACCTCTGCTCTTTGCTGAGAGCCGACGCTGTGATCTTCGAGGTACTCTGTGCGG 180
Qy 138 LeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAlaAla 157
Db 181 CTGAGGTGAAGAGACTTGGCTCAAGGAGGTGAAGTGAATGACACACAGGAGCT 240
Qy 158 LeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAspArg 177

```

Db 241 CTCCTGATTTACACATTTGCAATCTGAGATTGGGGATTTTGATGAGCCCTTGACAG 300
 QY 178 GlnH1LeuAlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIleVal 197
 Db 301 GAGCATTACCAAAATATATATACATCTCAGACAGACGACTTAGAGCAAAATCGTG 360
 QY 198 GluPheH1eH1AsnH1sIleGlyGlnThrProAlaGluSerAspPheGlnLeuLeuGlu 217
 Db 361 GAATTCACCACTAACCACTTGACAAACACGACAGAAATCAGATTTCCAGCTCTAGAG 420
 QY 218 IleAlaArgArgLeuGluMetTyrGlyIleArgLeuH1sProAlaLysAspArgGluGly 237
 Db 421 ATGGCCGTCGCTGAGATGATGAGATCCGGTTCGACCCGGCCAAAGGACAGAGAGGC 480
 QY 238 ThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLys 257
 Db 481 ACBAATATCAATCTGGCCGCTTGCCACACGCGAAATCTAGTGTTCAGGGTTTCATAG 540
 QY 258 IleAsnAlaPheAsnThrAlaLysValArgLysLeuSerPheLysValArgPheLeu 277
 Db 541 ATCAATGCCCTTCAACTGGGCGCAAGTCCGAGAGCTGAGCTTCAAGAGAGAGCGCTTCTC 600
 QY 278 IleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAla 297
 Db 601 ATCAAGCTCGGCGCAGATGCAATAGTCGTACCAAGATACCTTGAAATCTTGATGGCC 660
 QY 298 SerArgAspPheCysLysSerPheThrLysIleCysValGlnH1eH1sAlaPhePheArg 317
 Db 661 AGTCGGGATTTCTGCAGATCTCTTGAAATCTGTGTGAACATCATGCTTCTTTAGA 720
 QY 318 LeuPheGlnGluProLysProLysProLysProValLeuPheSerArgLysSerPhe 337
 Db 721 CTTTGAAGAGCCCAACCAAGCCCAAGCCGCTCTTGTAGCCGGGGGTGATCATTT 780
 QY 338 ArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrAlaLysGluGlyIleLys 357
 Db 781 CGGTTCAGTGTGGACTCAGAAAGCAGGTTCTGATGTGTAAAGAGAGAGCAATAG 840
 QY 358 LysValGlnPheGlnLysArgLysH1sSerLysIleH1sSerIleArgSerLeuAlaSerGln 377
 Db 841 MAGGTGAGTTTAAAGAGACACAGCAATTCATTCTGAGACCTTGCTTCACAG 900
 QY 378 ProThrGluLeuAsnSerGluValLeuGlnSerGlnInsThrSerLeuThrPhe 397
 Db 901 CCAACAGAACTGAATTCGAAAGTGTGAGACATCTCAGACAGACACAGCCCTTACATT 960
 QY 398 GlyGlnGlyAlaLysSerProGlyGlyGlnSerCysArgArgGlyLysGluProLysVal 417
 Db 961 GAGAAAGGTGCCAATCTCCAGGAGGCGCAGAGCTGCCGCGAGAGAAAGGAAAGGTT 1020
 QY 418 SerAlaGlyLysProGlySerH1sProSerProAlaProArgArgSerProAlaGlyAsn 437
 Db 1021 TCCGCCGGGGAGCGGGGTGCAACCGAGCTCTGCCGCGAGAGAGAGCCCGGGGTAAAC 1080
 QY 438 LysGlnAlaAspGlyAlaAlaSerAlaProThrGlnGluGluGluValValLysAsp 457
 Db 1081 AACCAAGCGGAGCGAGCGCTCGGCGCCACAGAGAGAGAGAGAGAGGTCTTAAGGAT 1140
 QY 458 ArgThrGlnGlnSerLysProGlnProProGlnProSerThr----- 471
 Db 1141 AGGACCCAGAGAGTAACTCAGCCCCGACGCAAGACAGAGTCCAGCATCCCGGGCT 1200
 QY 471 ----- 471
 Db 1201 GCCAGAGCGACGACTTCATTCTTTGATGACTGACGTGACGTAGATAGCAATAC 1260
 QY 471 ----- 471
 Db 1261 GACCTTCTTAGAGAAAGAGCGTCCGATCCGGTCCGATCCGGTCCCGATCCCAACA 1320
 QY 472 -----GlySerLeuThrGlySerProH1sLeuSerGlyLysSerValAsn 486
 Db 1321 AATGACATTATAGAGGCTCCCTGACTGGACATCTCACCCTTCCAGAGCTGTCTGTGAC 1380

QY 487 SerGlnGlyValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAspThr 506
 Db 1381 TCGACAGGGGGAGGTGGCCCCCTGCAACGTGACTTGTCTCCAACTGAGAGCCGACACC 1440
 QY 507 LysGlnAlaSerProLeuIleSerProLeuLeuAsnAspGlnAlaCysProArgThrAsp 526
 Db 1441 AAGCAGGCTCTCCCTTGATCAGCCCGCTGCGAAAGACAGAGGCTCCCGCGAGAGGAC 1500
 QY 527 AspGluAspGluGlyArgArgLysArgPheProThrAspLysAlaThrPheIleAlaLys 546
 Db 1501 GATGAGATGAGGGCGCGAGAGAGATTCACCACTGATTAAGCGATCTCATAGCTTAG 1560
 QY 547 GluValSerThrThrGluArgThrTyrLeuLysAspLeuGluValIleThrSerThrPhe 566
 Db 1561 GAAGTGTCAACACGAGGAAATCTTGAGAGATCTGAAATTCATCTTGCTGTGTTT 1620
 QY 567 GlnSerThrValSerLysGluAspAlaMetProGluAlaLeuLysSerLeuIlePhePro 586
 Db 1621 CAGAGCACAGTAGAGCAAAAGAGACCGCCATGCCGAGACCTGMAAGTCTCATATTCGG 1680
 QY 587 AsnPheGluProLeuH1sLysPheH1sThrAsnPheLeuLysGluIleGluGlnArgLeu 606
 Db 1681 AATTTGAACCTTTGACAAATTTCTACTTAATTTCTCAAGGAAATTGACACACACTT 1740
 QY 607 AlaLeuTrpGluLysArgSerAsnAlaGlnIleArgAspTyrGlnArgIleGlyAspVal 626
 Db 1741 GCCCTG----- 1746
 QY 627 MetLeuLysAsnIleGlnGlyMetLysH1sLeuAlaH1sLeuTrpLysH1sSerGlu 646
 Db 1746 ----- 1746
 QY 647 AlaLeuGluAlaLeuGluLysGlyIleLysSerSerArgArgLeuGluLysPheCysArg 666
 Db 1747 -----TGC--- 1749
 QY 667 AspPheGluLeuGlnLysValCysTyrLeuProLeuAsnThrPheLeuLeuArgProLeu 686
 Db 1750 ---GTTGAGCTGACAGCTGCTAGGGGAGAGACCCCTGCTGCTCTT----- 1791
 QY 687 H1sArgLeuMetH1sTyrLysGlnValLeuGluLysLeuCysLysH1sH1sProProSer 706
 Db 1792 -----CCGCTGCA 1800
 QY 707 H1sAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIleThrGluMetValAlaGln 726
 Db 1801 GACTGTATGTCTTCTTCCACAGCCGCTTTGGCAGAGATCAGAGATGTGTGGCAGAG 1860
 QY 727 LeuH1sGlyThrMetIleLysMetGluAsnPheGlnLysLeuH1sGluLeuLysLysAsp 746
 Db 1861 CTCACGATGATGATCAAGATGAGAAATTCAGAAACTGACAGAACTCAAGAAAGAT 1920
 QY 747 LeuIleGlyIleAspAsnLeuValProGlyArgGluPheIleArgLeuGlySerLeu 766
 Db 1921 TTGATTTGGATTAACATCTGTGTGTTCCGGAAGGAGATCATCTGCTGGCAGCCCTC 1980
 QY 767 SerLysLeuSerGlyLysGlyLeuGlnGlnArgMetPhePheLeu----- 781
 Db 1981 AGCAGCTCTCGGGAGAGGGGCTCCAGCAGCCGATTTCTTCTGTGATGAGAGAGAGC 2040
 QY 781 ----- 781
 Db 2041 GGCTTGTCTCAAGAGATTGTGTCACTGGGCAAGCAAGCTCCCAAGGCCGAGCTCG 2100
 QY 782 -----PheAsnAspValLeuLeuLysTyrThrSerArgGlyLeuThrAlaSerAsnGln 798
 Db 2101 GCCACCAATTAAACACACCTCTGTATACAGAGCGGGGGCTGACGCGCTTCATATCAG 2160
 QY 799 PheLysValH1sGlyLysLeuProLeuTyrGlyMetThrIleGluLysSerGluAspGlu 818
 Db 2161 TTTAAAGTCCACGGGAGCTCCCGCTCTATGGCATGACGATTGAGAGAGACGAGACGAG 2220

```

Qy      819  TrrgYValProHsCYaleuThrluAryglYlnAryglInserIleIleValAlaAla 838
Db      2221 TGGGGGGGTGCCCACTGCTTACCTCCGGGGCCAGCCGCTCATCATCATGCGCCG 2280
Qy      839  SerSerArgSerGluMetGluYleTrrValGluAplIlegImetAlaIleAspLeuAla 858
Db      2281 AGTTCCTCGTCCGAGATGAGAGAGTGGGTGAGGACATCCAGATGCGCATTGACCTGGCG 2340
Qy      859  GlulYSerSerSerProAlaProGluPhleuAlaSerSerProProAlaAsnYleSer 878
Db      2341 GAGAAAGACACACACCCCGCCCTGAGATTCTTGCCAGACACCCCTCGACAAACAGTCC 2400
Qy      879  ProAspGluAlaTrrAlaAlaAspGlnYleSerGluAAspAAspLeuSerAlaSerArgThr 898
Db      2401 CCTATGAAAGCCACCGCGCTGACACAGAGTCAAGAGATGACTGAGCGCTCCGCGACA 2460
Qy      899  SerLeuGluArgGlnAlaProHsAArgYleAanthMetValHsValCYstrPhsArg 918
Db      2461 TCGCTGAGGCGCCAGGCGCCGACCGCGGCAACAATAGTGCAGTGTCTGGCACCGC 2520
Qy      919  AanthSerValSerMetValAspPheSerIle-----AlaValGlu 932
Db      2521 AACCCAGCGCTCTCATGTGTGACTTCAGCATTCGCACTGAGGCGCTCTTCCCTTCAG 2580
Qy      933  AasnGlnYleSerGlyAanYleuAryglYAspPheYAsnSerAasnGlyTrrGlnYleu 952
Db      2581 AATGATGTCCTGGAACCTGCTGAGGAATTCAAACACACACCGGTGCGACAGCTG 2640
Qy      953  TrrValAlaPheTrrAanPheCYleuPhePheTrrYlySerHsIleGlnAspAsnHsIlePro 972
Db      2641 TGGGTGTGTTCAAACTTCTGCTGTGTTCTTTCACAAATCACACCGAGAACATCATCTCC 2700
Qy      973  LeuAlaSerLeuProLeuLeuGlyTrrSerLeuThrIleProSerGluSerGluAsnIle 992
Db      2701 CTGCGCCAGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
Qy      993  GlnYAspTrrValAlaPheYleuHsIlePheYSerHsIleValTrrYtrPheArgAlaGlu 1012
Db      2761 CAGAAAGACTRACGTTCACACTGCACTTCAAGTCCACGCTCACTACTACTAGGCGGAA 2820
Qy      1013  SerGluTrrYtrPheGlu 1018
Db      2821 AGCGAGTRACAGCTTGGAA 2838

RESULT 8
BC030329 3902 bp mRNA linear ROD 21-OCT-2003
LOCUS Mus musculus FERM, RhogEF (Arhgef) and plecterin domain protein 1
DEFINITION (chondrocyte-derived), mRNA (cDNA clone IMAGE:5376197), partial
cds.
ACCESSION BC030329
VERSION BC030329.1 GI:20987935
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus (house mouse)

REFERENCE
AUTHORS Strausberg, R.D., Collins, F.S., Wagner, L.H., Derge, J.G.,
Klauser, R.D., Collins, F.S., Wagner, L.H., Schmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F.,
Dachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Ubedi, T.B., Toohily, S.,
Carninci, P., Prange, C., Raha, S.S., Loughlin, N.A., Peters, G.J.,
Abramson, R.D., Muller, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Wojcik, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

```

```

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
Boufard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Scherer, A., Schein, J.E., Jones, S.V., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3902)
Strausberg, R.
Direct Submission
Submitted (07-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Ahter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Boufard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Maduro, Q.L., Maiello, C., Maskeri, B., Masriani, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W.,
Turgeson, C., Vogt, J.L., Walker, M.A., Weatherly, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ULNL at: http://image.llnl.gov
Series: IRAK Plate: 65 Row: m Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
Location/Qualifiers
1. 3902
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5376197"
/tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month
old mouse. Taken by biopsy."
/clone_1lb="NCI_CGAP_Mam2"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
<1. 3902
/gene="Farpl"
/note="synonym: Cdep"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
<1. 2510
/gene="Farpl"
/codon_start=3
/product="Farpl protein"
/protein_id="AAH0329.1"
/db_xref="GI:20987936"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
/translaction="TRPEVARRLEMTGIRLHPAKDBRGTKINLAVANTGILVFGFTK
INLFWNAKVRKLSFKRRLIKRDPVNSYQDTLIEIAGRGDCEKSFKICVHNHF
PRLEPRKPKPVLRVLSRSGSSFRSGRTQOVDYKSGGKKVQPERKSKHNSTRS
LVSQPGVADNSVVKQSPQASLTFGEGBESPGQSCQOAKETACTLIEGPHOSPALP
KSPGSPVADNSVVKQSPQASLTFGEGBESPGQSCQOAKETACTLIEGPHOSPALP
OGGAAPANVTLSPNLSPDNKOASPLISPLINDOACPTDDEBGRARRRPTDRAVYA
KEVSTERTYTLKOLEVIASWFGSTVSKEDSMPALKSLIRPNPEPLHKFPTNFKETE

```

ORLALWEGNSNAHVGRDRIQIDVMLKNIQGMKHLAHLWKHSEALEAETSIKGSR
LEHFCDPELOKVCYLPLNTFLPLRPLRLMHWYGLVERLCKHP PNHADFRDRAALA
EIRWVAOLAHGWTIMKMNFOKHELKDLGINDLVTPGEPFRLSLSLGSKGIO
RMPELNDVLVTSRGLTASNORKGOLPLYMTI ESEEGVGHCLTRORSI
IYAASSREWEKEMEDIQMAIDLAESKNGTPPELASSPPDNKSPDEATRADESED
LSASRTSLERQAPHRGNTVWVCWHRNSTVSWDFSI AVENQISGMLRKFKNSEMO
KLMVVTNCTLFYKSHODSHPLASLPLLOYSLTIPSESNHKDYVFKLHFKSHYV
FRASEYTERFERNMEIVIRSATSSASRAHILSHKESHLY"

misc_feature

/gene="Parp1"

/note="RhoGEF; Region: Guanine nucleotide exchange factor
for Rho/Rac/Cdc42-like GTPases"

misc_feature

/db_xref="CDD:smat00325"

1695..1946

/gene="Parp1"

/note="PH; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntrrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinaemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"

/db_xref="CDD:smat00233"

misc_feature

2175..2465

/gene="Parp1"

/note="PH; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntrrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinaemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"

/db_xref="CDD:smat00233"

ORIGIN

Alignment Scores:

Pred. No.: 9,39e-316 Length: 3902
Score: 3963.50 Matches: 758
Percent Similarity: 93.8% Conservative: 22
Best Local Similarity: 91.1% Mismatches: 49
Query Match: 72.6% Indels: 3
DB: 6 Gaps: 2

US-09-555-342B-2 (1-1045) x BC030329 (1-3902)

QY 217 GltuilealaargrgleuglumeTYrGlyleargreuhisProAlaIysAspArgLu 216
|:::|
Db 12 GAGGTGGCCCGTGGCTAGAGATGTAACGGATCCGTTACACCTGCGAAGCAAGGAA 71
QY 237 GlyThrLysleleuLeuAlaValAlaAsnThrGlylleuValPheGlnGlyPheThr 256
|:::|
Db 72 GGCACCAAGATCAACCTGCTGCGCAACACGTAATTCAGTGTTCAGGCTTCACG 131
QY 257 LysleleuAsnAlaPheAsnTPAlaIysValArgLysLeuSerPheLysAspGlyAspArgPhe 276
|:::|
Db 132 AAATCAATGCTTTTAACTGGCCCAAGGTGCGAAGCTTCAAGCCGAAGCCGCTTT 191
QY 277 LeuileLysLeuArgProAspAlaAsnSerAlaTyrcGlnAspThrleuGlnPheLeuMet 296
|:::|
Db 192 CTTATCAAGCTGACCTGACGTGAATGCTCATACCAAGACACCTTGAGGATTCCTGATG 251
QY 297 AlaSerArgAspPheCysLysSerPheTrpLysleCysValGluHisAlaPhePhe 316
|:::|

Db 252 GCCGGCCGAGACTTCGAAAGCTTTCTGAAAAATATGTGTGAGACCAATGCTTCTTC 311
QY ArgLeuPheGlnGluProLysProLysProLysProValleuPheSerArgGlySerSer 336
317 |:::|
Db 312 AGACTTTTGAGAACTTAAGCCAAAGCCAGTCTTCTTCAGCGGAGGTGCTCC 371
QY PheArgPheSerGlyArgThrGlnLysGlnValleuAspTyrValLysGlnGlyLysHis 356
337 |:::|
Db 372 TTCGGTTTATGTGTGCACTCGAAGCAAGTCTTGATATATGTAAAGAGCGGCGCAC 431
QY 357 LysLysValGlnPheGlnArgLysHisSerLysleHisSerleleArgSerLeuAlaSer 376
|:::|
Db 432 AAAGAAGTCAAGTTTAAAGAAAGCAACACAGATCCACTCCACCGAGGCTGTCTCA 491
QY 377 GlnProThrGluLeuAsnSerGluValleuGlnGlnSerGlnGlnSerThrSerLeuThr 396
|:::|
Db 492 CAGCCCAAGACACCGAATTCAGAAAGTCCAAAACAGTCTCCACAGGCGCAGACCTTAAG 551
QY 397 PheGlyGlnGlyAlaGluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLys 416
552 |:::|
Db 552 TTTGAGAAAGGTACGAGTCTCGGGGTGGCCAGAGCTGCAGCAAGCAAGAAACAAA 611
QY 417 ValSerAlaGlyGluProGlySerHisProSerProAlaProArgSerProAlaGly 436
612 |:::|
Db 612 GCCTGTACCTTGAGAGTTGGGACCTCAACAAAGCCCGCTCCCTAAGAGCCCTCCAGGG 671
QY 437 AsnLysGlnAlaAspGlyAlaAlaSerAlaProThrGlnGlnGlnGlnValVal--- 455
|:::|
Db 672 AGCAAAAGCTGCAGATGAAACCAACCGTGGCCGCTGAGAGAAAGAGAAAGAGGT 731
QY 456 ---LysAspArgThrGlnGlnSerLysProGlnProGlnProGlnProLysSerLeu 474
732 |:::|
Db 732 GGCAGACAGGACCTCGGCAAGTAACCTCAGAGCCGCCAGCAAGCAAGGCTCCCTG 791
QY 475 ThrGlySerProHisLeuSerGluLeuSerValAsnSerGlnGlyValAlaProAla 494
Db 792 ACTGTAAGCCCTCACTTCAAGAGCTGCTCATCACTCAAGAGAGAGAGCGGCCAGCC 851
QY 495 AsnValThrLeuSerProAsnLeuSerProAspThrLysGlnAlaSerProLeuLysSer 514
852 |:::|
Db 852 AACGTAACTGTCTTCTCCAACTGAGTCTTGACAAACAGAGGCTCTCCCTGATCAGC 911
QY 515 ProLeuLeuAsnAspGlnAlaCysProArgThrAspAspGluAspGluArgArgLys 534
912 |:::|
Db 912 CCGCTCTGTAAGACCAAGCAAGCTGCGCAAGACCAAGCAAGAAAGAGGCGCGCAAG 971
QY 535 ArgPheProThrAspLysAlaTyrPheleleAlaLysGluValSerThrThrGluArgThr 554
|:::|
Db 972 AGATTTCCAACGACAAAGCGTACTACATTTGCTAAGAAAGTCTCCACACAGAGCCACG 1031
QY 555 TyrLeuLysAspLeuGluValIleThrSerTrpPheGlnSerThrValSerLysGluAsp 574
1032 |:::|
Db 1032 TACCTAAGAGACTTGAAGTCAATGCTTCTGCTTCAAGACCAATCAAGAAAGAGAC 1091
QY 575 AlaMetProGlnAlaLeuLysSerLeuIlePheProAsnPheGluProLeuHisLysPhe 594
|:::|
Db 1092 TTCATCCCGGAAGCCTTGAAGAGTCTCATATTCACAAATTTTGAACCTTTGACAAAGTTTC 1151
QY 595 HisThrAsnPheLeuLysGluIleGlnGlnAlaGluAlaLeuTrpGluValArgSerAsn 614
Db 1152 CATACCAATTTTCTGAAGAAATTTGAGCAACGACTTGCCTGTGGAGAGGCCCTCAAT 1211
QY 615 AlaGlnIleArg---AspTyrGlnArgIleGlyAspValMetLeuLysAsnIleGlnGly 633
|:::|
Db 1212 GCCCAAGTCAAGAGAGATTAACAAAGATCGGGAGCTCATGCTGAAAAATTCAGGCGC 1271
QY 634 MetLysHisLeuAlaAlaHisLeuTrpLysHisSerGlnAlaLeuGluAlaLeuGluAsn 653
Db 1272 ATGAAGCACTGGCGGCGCCACCTGTGAGAGCAACAGTGAAGGCGCTTGAGGCGCTGAGACC 1331
QY 654 GlyIleLysSerSerArgArgLeuGluAsnPheCysArgAspPheLysGlnLysVal 673
|:::|
Db 1332 TCCATCAAGAGCTCGGGGCGGTGGAACATTTCTGCGAGACTTCGAGGCTGCAAGAGGTG 1391

QY 674 CyeTyrLeuProLeuAenThrPheLeuLysrProLeuHisArgLeuMetHisTyrLys 693
 Db 1392 TGTAACTGGCGGCTCAACACTTCTCTGCGCGCGCTGCAAGCTCATGCACTAAAG 1451
 QY 694 GlnValLeuGluArgLeuCyValPheHisIleProSerHisIleAspPheArgAspCys 713
 Db 1452 CATCTCTGGAGAGGCTGTGCAACACCAACCAACCGCGCACTTCAGGAGACTGC 1511
 QY 714 ArgAlaAlaLeuAlaGluIleThrGluMetValAlaGlnLeuHisGlyThrMetIleLys 733
 Db 1512 AGAGCTCGCTGGCGGAGATACAGAGATGCTGGCACAGCTGATGATCATATCAAG 1571
 QY 734 MetGluAsnPheGlnLysLeuHisGluLeuLysValAspLeuIleGlyIleAspAsnLeu 753
 Db 1572 ATGGAGAACTTCCAGAAAGCTGCATAGCTCAAGAAAGTCTGATGGCATTAACAATCTT 1631
 QY 754 ValValProGluArgGluPheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGly 773
 Db 1632 GTGACCCCAAGAAAGGAGTTCATCCGCTGGCGAGCTTCAGCAAGCTTCGCGGAAGGG 1691
 QY 774 LeuGlnGlnArgMetPhePheLeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeu 793
 Db 1692 CTTCAAGAGCGCATGTCTTCTCTGTCACAGATGCTGTGATACAGCCGCGGAGACTG 1751
 QY 794 ThrAlaSerLeuGlnPheLysValHisGlyGlnLeuProLeuTyrGlyMetThrIleGlu 813
 Db 1752 ACCGCACTTAATCAAGTTAAAGTCCAGAGACGCTCCACCTTAAGCATGACATGAG 1811
 QY 814 GluSerGluAspGluTyrGlyValProHisCysLeuThrLeuArgGlyGlnArgGlnSer 833
 Db 1812 GAGAGTAGAGAGAGAGTGGGGTGTCCCACTGTCAGCCCTGGCGGCGCAGCGCAATCT 1871
 QY 834 IleIleValAlaAlaSerSerArgSerGluMetGluLysTyrValGluAspIleGlnMet 853
 Db 1872 ATCATCTGCTGCTGCGCACTCCCGCTCAGAGATGAGAAATGAGATGAGAGCATCAGATG 1931
 QY 854 AlaIleAspLeuAlaGluLysSerSerSerProAlaProGluPheLeuAlaSerSerPro 873
 Db 1932 GCCATTGACCTGGAGAGAAAGCAATGAGACCCCGCGAGCTCTTGGCAGAGAGCCC 1991
 QY 874 ProAspAsnLysSerProAspGluAlaThrAlaAlaAspGlnLysSerGluAspAspLeu 893
 Db 1992 CTTGACAAAGTCCCGGATGAGACCAAGCGGCTGACCGAGGTGAGAGATGACTG 2051
 QY 894 SerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgGlyAsnThrMetValHis 913
 Db 2052 AGTGCTCTCGACATCGCTGAGCGCGAGCCCTCAGCGTGCAACAGATGCTGCAC 2111
 QY 914 ValCysThrPheIleArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
 Db 2112 GTGTGCTGGATCGGACACCAAGTGTCTCCATGTGATTCATTCATTCGCTGTGAGAAC 2171
 QY 934 GlnLeuSerGlyAsnLeuLeuArgLysPheLysAsnSerAsnGlyTyrTrpGlnLysLeuTyr 953
 Db 2172 CAGCTCTCTGGAACTTGGCTGAGAAAGTTCAGAAACAGCAATGCTGGCAGAACTGTGG 2231
 QY 954 ValValPheThrAsnPheCysLeuPhePheTyrLysSerHisGlnAspAsnHisIleProLeu 973
 Db 2232 GTGGATTACCAAACTTCTGCTGTCTTCTCAAAATCACACGAGACAGTCAACCCCTT 2291
 QY 974 AlaSerLeuProLeuLeuGlyTyrSerLeuThrIleProSerGlnLysSerGluAsnIleGln 993
 Db 2292 GCCAGCTGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2351
 QY 994 LysAspTyrValPheLysLeuHisPheLysSerHisValTyrTyrPheAlaGluSer 1013
 Db 2352 AAGGACATATGTTCAGAGCTCATTTCAAGTCCACAGTATCTTCAAGGCTGAGAAC 2411
 QY 1014 GluTyrThrPheGluArgTyrPheMetGluValIleArgSerAlaThrSerSerAlaSerArg 1033
 Db 2412 GAATACACATTTGAAAGGTGATGAGAGTATCCGAAGTGCACACACTGAGCTCTCGG 2471

QY 1034 ProHisValLeuSerHisLysGluSerLeuValTyr 1045
 Db 2472 GCCCAGATCTTGAATCAAAAGAAATTCACCTGTAC 2507
 RESULT 9
 LOCUS CO895396
 DEFINITION Sequence 16 from Patent EP1471153.
 ACCESSION CO895396
 VERSION CO895396.1 GI:55467577
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE
 1
 AUTHORS Certa, U., Poser, S. and Weyer, K.
 TITLE Transcriptional activity assay
 JOURNAL Patent: EP 1471153-A 16 27-OCT-2004;
 F. HOFMANN-LA ROCHE AG (CH)
 FEATURES
 source
 1. 3997
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,42e-231 Length: 3997
 Score: 2941.50 Matches: 580
 Percent Similarity: 70.0% Conservative: 162
 Best Local Similarity: 54.7% Mismatches: 261
 Query Match: 53.8% Indels: 57
 Gaps: 14
 DB: 2
 US-09-555-342B-2 (1-1045) x CO895396 (1-3997)
 QY 1 MetGlyGluIleGluGlnArg-----ProThrProGlySerAlaGluValAla 16
 Db 118 ATGGGGAGATAGAGAGAAACATACAGAGTCTGACAGCTCAGAGGATGCGTTGGGTGCC 177
 QY 17 ProGluAsnSerGlyIleSerThrLeuGluArgGlyGlnLysProProThrProSer 36
 Db 178 CAGACCCCTGTGGAGATGACCTTGACCTGGGAGACTCTCTTCCCAAGATGCAA 237
 QY 37 GlyLysLeuValSerIleLysIleGlnMetLeuAspAspThrGlnGluAlaPheGluVal 56
 Db 238 GAGAAAGACCTGACACCTCAGAGTAAGCTGTCGACAAACCAATGAAATATTTGACATT 297
 QY 57 ProGlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHisIleAsnLeuVal 76
 Db 298 GAGCTTAAATGCGATGGCCAGATTAATCTGACACAACTGTGAAAGCTTTAAACCTGTA 357
 QY 77 GluGlyAspTyrPheGlyLeuGluPheProAspHisLysLysIleThrValITrPLeuAsp 96
 Db 358 GAATGATCACTACTTGGGATGAGGTTCAAAATTAATCAAGTCTTACTGATTTGGCTTGA 417
 QY 97 LeuLeuLysProIleValLysGlnIleArgArgProLysHisValValLysPheVal 116
 Db 418 CCAATGAAACCATCATTAAGCAAAATAGCAAGGCAAAAGATGTGCTGCTGCTGCTGCTG 477
 QY 117 ValLysPhePheProPheAspHisThrGlnLeuGlnGluGluLeuThrArgTyrLeuPhe 136
 Db 478 GTAAATTTTTCACCTGATCTCTGTCACTCAAAAGAAATTAACAAAGATTAATCTTGT 537
 QY 137 AlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAla 156
 Db 538 GCTTGAACATTAAGAGAGACTCTGCTGGAAGAGCTTTGACTGTGTGACACACAGCGG 597
 QY 157 AlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAsp 176
 Db 598 GCCCTTCAAGTCCCATCTCTGCAAGTCCGAAATAGAGATTAAGATGAAGACCTTGAC 657

QY 177 ArgGluHisLeuAlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIle 196
 Db 658 GCGAGCAGCACTTCAAGATGAGATGATTTGCTGGCCAGCAGCATGCTTGGAGAGATA 717
 QY 197 ValGluPheHisValAsnHisValIleGluInThrProAlaGluSerAspPheGlnLeu 216
 Db 718 CTGAATTCATCATGAAAGCAGTGGGCGAGACCTGCTGATGCGATTTCCAGGTCTC 777
 QY 217 GluIleAlaArgArgLeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGlu 236
 Db 778 GAATGCTCGAAAGTGGAAATGTACGCGATCATGATTTCACTGCTTGCAGAGGAA 837
 QY 237 GluThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
 Db 838 GGAACCAAGATTCACCTGCGCATTTCCACATGGGTCTACTGCTTCCAGGCGACACC 897
 QY 257 LysIleAsnAlaPheAsnTrpAlaLysValArgLysLeuSerPheLysArgLysArgPhe 276
 Db 898 AAATCAACACTTCACTGATGCTCCAAAGTCCGTAATTAAGCTTCAAGAGAAAAGATT 957
 QY 277 LeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
 Db 958 CTATCAAACTTCATCCAGAGGTTCTATGACCTTACAGAGACATTTAGAAATTTGTG 1017
 QY 297 AlaSerArgAspPheCysLysSerPheTrpLysIleCysValGluHisValaPhePhe 316
 Db 1018 GGTAGTGAATGATGATGATGAACTTGGAGATTTGTGTGAGATCACTACCTTTT 1077
 QY 317 ArgLeuPheGluGluProLysProLysProLysProValLeuPheSerArgLysSer 336
 Db 1078 AGACTTTGGACCACTTAAGCCAAAGCAAAAGCCGCTTCTTCAAGCCGGGCTCCTCC 1137
 QY 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyHis 356
 Db 1138 TTGAGATACAGTGGAGAACTCAGAAACAATGATATTATTCAAGCACTGGATG 1197
 QY 357 LysLysValGlnPheGluArgLysHisSerLysIleHis---SerIleAspSerLeu--- 374
 Db 1198 AAGAGATTCATATGAAAGAGAGCAGCAGAACCCACACGTCCTGTCAGACTGACT 1257
 QY 375 AlaSerGlnProThrGluLeuAsnSerGluValLeuGluInSerGlnInSerThrSer 394
 Db 1258 GCAGACCTACCAAAACAG-----AGC 1278
 QY 395 LeuThrPheGlyGluGlyAlaGluSerProGlyGlyInSerCysArgArg----- 411
 Db 1279 ATCTCATTTCCCGAGGATTTGAGGACTGCTGCCATCTTCAAGCAATGCTTTTAC 1338
 QY 412 -----GlyLysGluProLysValSerAlaGlyGluProGly----- 423
 Db 1339 TCGCTCTCTCCCTCCACTCTGCTCCCTCTGCTGCTGCAGAGTTTAAAGCAGCAGCAGC 1398
 QY 424 -----SerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAlaAsp 441
 Db 1399 TCCCTACAGATCCCGAGGTTCTTCACTGATCAAGAGTCACTGACAGAGGCCAGTGA 1458
 QY 442 GlyAlaAlaSerAlaProThrGluGluGluGluValValLysAspArgThrGlnGln 461
 Db 1459 GCAGTGGCTGAGGCCCCGAC-----ACACCATCG 1488
 QY 462 SerLysProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
 Db 1489 GCCACGCTCTGGGGCCCCCGCAGCTCCAGCTGCTGTCAGGCTTTCCAGAGAGCTCT 1548
 QY 479 HisLeuSerGluLeuSerValaAsnSerGlnGlyValAlaProAlaAsnValThrLeu 498
 Db 1549 CAGCTTCTCTCCCTCCAGCGGAGAGGCCCTGAGTGTGAGCCCTGCA-----TTT 1599
 QY 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuLysSerProLeuLeuAsn 518
 Db 1600 CAGGTGCTTTGGGCGCAGCTGAACAGGCTCATCCCACTCTGAGCCCTGTCTCAGT 1659

QY 519 AspGlnAlaCysProArgThrAspArgGluAspGluGlyValArgArgLysArgPheProThr 538
 Db 1660 GATGCTGGGGGAGCGCGGANTGACTCGAG---GAGCCAGAGACAAAGCGCGCTGCA 1716
 QY 539 AspLysAlaTyrPheIleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAsp 558
 Db 1717 GACGAGGCTTACTTCTTACATCAAGATTCCTCGCTACAGAAAGCAATACCTCAAGAT 1776
 QY 559 LeuGluValIlePheSerTrpPheGlnSerThrValSerLysGluAspAlaMetProGlu 578
 Db 1777 TTGAAGTTATTAACCTGTGGTTCCGACGCCAGTGTGAAGAGAGCCGATGCTTGGC 1836
 QY 579 AlaLeuLysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAspPhe 598
 Db 1837 ACTGTATGACGCTGCTTCTTCCATCATGATCCCATGTATGATGATTCACAGAGCTTC 1896
 QY 599 LeuLysGluIleGluGlnArgLeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIleArg 618
 Db 1897 CTGCGGAGGTGGAGCAGAGGCTGCGCACTTGGGAAGGCCCTCCAAAGCCACACAAA 1956
 QY 619 ---AspTyrGluArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLeu 637
 Db 1957 GCGATGATCAACGAATCGGGACATCTGCTCAGGAACATGCGCCATTAAGAGACTTT 2016
 QY 638 AlaAlaHisLeuTrpLysHisSerGluAlaLeuGluAlaLeuGluAsnGlyLysSer 657
 Db 2017 ACCAGCTACTTCCAAAGACATGACGAGTCCCTTAACAGAACTGGAAAAGCTACCAAGC 2076
 QY 658 SerArgArgLeuGluAsnPheCysArgAspPheGluGlnLysValCysTyrLeuPro 677
 Db 2077 TGTAAAGATTTGAGGACAGTGTACAGAGATTGACCTCAGAAAGTCTGCTACTTGGCT 2136
 QY 678 LeuAsnThrPheLeuLeuTrpProLeuHisValArgLeuMetHisTyrTyrGlnValLeuGlu 697
 Db 2137 CTCAACAGCTTCTCTGTAAAGCCATCCAGCGGCTCTGTGACTACCGCTGCTGCGC 2196
 QY 698 ArgLeuCysLysHisValProProSerHisValaAspPheArgAspCysArgAlaAlaLeu 717
 Db 2197 CGCTATGGGACATTAACAGCCCGGCGACCATGACTAGCAGTGCATGCCATGACGCCCTG 2256
 QY 718 AlaGluIleThrGluMetCysValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPhe 737
 Db 2257 AAGCCATCAGAGAGGTGACCAACACATCAAGCACATTTCTCATCCGCTGAGAACTTG 2316
 QY 738 GlnLysLeuHisGlyLeuLysLysAspLeuIleGlyIleAspPheLeuValValProGly 757
 Db 2317 CAGAGCTTAACGAGCTGCGAGCGGACCTGGTGGCATGAGAACTCATGCTCTGCTGGC 2376
 QY 758 ArgGluPheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArg 777
 Db 2377 AGGAGTTTCATCCGTGAGGGCTGCTTCAAGCTCACCAAGAAAGGCTGCGACAGAGG 2436
 QY 778 MetPhePheLeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsn 797
 Db 2437 ATGTTTTTCTGTTCACATATGTGTGTGTACACAAAGAAAGAAAGTTCAGAGGACAGC 2496
 QY 798 GlnPheLysValHisGlyGlnLeuProLeuTyrGlyMetThrIleGluGlnSerGluAsp 817
 Db 2497 CACTTCGGAATCGGGGCTCTCTCCCTCCAAAGCATGCTGTGTGAAGAAATGATTAAC 2556
 QY 818 GluTrpGlyValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAla 837
 Db 2557 GAGTGTCTGTTCACACATGTTTCACTACATTCAGCGGCTCAGAAAACAATCGTGTGCA 2616
 QY 838 AlaSerSerArgSerGluMetGlyLysTrpValGluAspIleGlnMetAlaIleAspLeu 857
 Db 2617 GCCAGCATCTCGGTGAGAAAGAAAGTGAATGCTGTGAACCTCCGAGATCCAAAGCA 2676
 QY 858 AlaGluLysSerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
 Db 2677 GCCAAGAGTGGCGGTACACGCGCTTGCATGCGCAGCGGCACATGTGTGCACTGCTCC 2736
 QY 874 ProAspAsnLysSerProAspGluAlaThrAlaAlaAspGlnGluSerGluAspLeu 893

```
Db 2737 CCC-----AGATCCCCCAAGAG---GTATCTGTGGAGCGAGGTCAAGAAATGATGCT 2787
Qy 894 SerAlaSerAlaThrSerLeuGluArgGlnAlaProHisValGlyAlaThrMetValHis 913
Db 2788 CCGGGGTCTCCGACACTCCCTGGAGGGGATGCCAGCACCCGGGCCAACACCAATGAC 2847
Qy 914 ValCysThrPheValSerGlnThrSerValSerMetValAlaPheSerIleAlaValGlnAsn 933
Db 2848 GTGTGCTGTGATCCGGAACACGAGGCTGTCCAGGCGAGACCAAGTGCAGCTGTGAGAAC 2907
Qy 934 GlnLeuSerGlyAlaLeuLeuArgGlyPheLeuSerHisSerAlaGlyTyrGlnLeuLeuTyr 953
Db 2908 CAGCTTCAGAGATATCTGCTAAAGAAAGTTCAAAAACATCATGCTGCGCAGAACTCTGG 2967
Qy 954 ValValIleThrIleAsnProGlyLeuPhePheTyrIleSerHisGlnAsnAsnIleProLeu 973
Db 2968 GTCTCTTTTACCACTCTGTGTTGTTCTTCTTACAAACTCATGATGATGATGATGATGATG 3027
Qy 974 AlaSerLeuProLeuLeuGlyTyrSerLeuThrIleProSerGluSerGluAlaGln 993
Db 3028 GCCAGCTCCCGCTGCTGGCTAGAGCTGAGCAATCCCGAGGAGCCGATGCGCATACAC 3087
Qy 994 LysAspTyrValPheLeuLeuIlePheLeuSerHisValTyrTyrPheArgAlaGluSer 1013
Db 3088 AAGACATATGTTTCAAGCTCAGATTCAATCCACGCTTACTCTTCCGGGCTGAGAGC 3147
Qy 1014 GluTyrThrPheGluArgTyrMetGluValIleArgSerAlaThrSerSerAlaSerArg 1033
Db 3148 AAGTACACATTTGAAGGTGATGATGAGGTGATCCAGGGGCGCACAGCTCAGCCGGAGG 3207

RESULT 10
AB018336 3997 bp mRNA linear PRI 10-JAN-2004
LOCUS Homo sapiens mRNA for KIAA0793 protein, partial cds.
ACCESSION AB018336
VERSION AB018336.1 GI:3882306
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Miyajima,N.,
Tianaka,A., Kotani,H., Nomura,N. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes. XI.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro
JOURNAL DNA Res. 5 (5), 277-286 (1998)
PUBMED 9872452
REFERENCE 2 (bases 1 to 3997)
AUTHORS Ohara,O., Suyama,M., Nagase,T., Ishikawa,K. and Kikuno,R.
TITLE Direct Submision
JOURNAL Submitted (08-OCT-1998) Otsu Ohara, Kazuo Research Institute,
Laboratory of DNA Technology; Yamta 1532-3, Kizuarzu, Chiba
222-0812, Japan (E-mail:cdna@infoc.kazuo.or.jp, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
FEATURES
source
1..3997
location/Qualifiers
1..3997
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hk05692"
/sex="male"
/tissue_type="brain"
/clone_id="obluescriptII SK plus"
/dev_stage="adult"
1..3997
/gene="KIAA0793"
<115..3282
/gene="KIAA0793"
/note="Start codon is not identified"
```

```
/codon_start=1
/product="KIAA0793 protein"
/protein_id="BA34513.2"
/db_xref="GI:40788367"
/translation="RMGEIGQYRVLTQAGRLGAQTPGVSTLEPGQTLLPRMQEKN
LHLRVKLDNTMWEI FDI EPCDQGVLTQWVKRLNVECDYFGMEPONTQSYMTL EPR
MKPILRIQRRKVVLLAVKFFPPDQGLQOEYTRTLPALQKRLDLERLTCADTT
AALLTSLHLSOSIGDYDETLREHLKXNEVLPQGOQLKELIEPHQKHOVOTPAESP
QVLEIARKMEVGRFPMASDRGTQGLASPKFMPKVCVETHTFFRLDDPKRAKAV
KRKRFILKLPVHNGPYQDTLEPLSLRDECKRPFKVCVETHTFFRLDDPKRAKAV
FESRGSFRYSGRKQKQVDFKDSGKRLPYERRRSKHTSVATLADLPKQISFP
EGRLTPASPSANAFYSLSPSTLVPSLPFPPKSSSLTDPQVSVVSPAAERSGAV
AGSPDPSAQLPPALQPPGLSTKSPQSPSSKSLSPALQVLPALAEQSSP
LSPVSLDAGAGMDCEBRHKVPADAEVFIYEKILATERTYLKDLVETPERSAV
VKEIDAMPATMTLLFSSNIDPIYEPHGRLEAVORKEILATMRGPKAKHTSGHRTGDL
LRMKQLKETSITFORHDEVLTLEKATKRCCKLEAVYKFELOKVCYDPLANTFLRP
IQRILHYRLRLRLCGHSFGNHDVADCHDLAKATEVTTTLQHLILRLNLQKLTBL
QRDLVGLNELLIARGREFIRGCILKCLKLQRMFLFSDMLLYTSKGVAGSHFRI
RGLLPLQGLMVEESDNEMSVPHCFETIYAAQTIYVAASTLEKEKMLDLSAIOAK
SGDDTPALPGRITVCTRPPSPNVEVLEQSSDDARGRSSLEGHGHRANTMVCW
YRNTSVSRADHSAAVENQLSGYLLRKTKSHGKQKMLVVFETNCLFXYKTHODDPIA
SLPULGYSVIFPEADGITHQVYFKLOFKSHVYFFRAESKTYTERMVEVIOGASSHG
RAPSIVDQGQPPSGLEGMVRKEE"
ORIGIN
Alignment Scores:
Pred. No.: 1,42e-231 Length: 3997
Score: 2941.50 Matches: 580
Percent Similarity: 70.0% Conserved: 162
Best Local Similarity: 54.7% Mismatches: 261
Query Match: 53.8% Indels: 57
DB: Gaps: 14
US-09-555-342B-2 (1-1045) x AB018336 (1-3997)
Qy 1 MetGlyGluIleGluGlnArg-----ProThrProGlySerArgLeuGlyAla 16
Db 118 ATGGGGAGATAGAGAGACATACAGAGTCTCTGACAGCTCAGAGGAGATGCCCTTGAGGCC 177
Qy 17 ProGluAsnSerGlyIleSerThrLeuGluArgGlyGlnIlyProProProThrProSer 36
Db 178 CAGACCCCTGTGGAGATTACACCCCTTGAGCTGGCGAGACTCTCTTCCCGCAGAAATGCAA 237
Qy 37 GlyIleLeuValSerIleLeuIleGlnMetLeuAspAspThrGlnGluAlaPheGluVal 56
Db 238 GAGAAAGCACTGCACTCAAGATTAAGCTCTGCAACAACCATGAGAAATATTGACATTT 297
Qy 57 ProGlnArgAlaProGlyIlyValLeuLeuAspAlaValCysAsnHisLeuAsnLeuVal 76
Db 298 GAGCTTAATGCCATGCGCAGATTTACTGACACAACTGTGGAAGCGTTTAAACCTGTGTA 357
Qy 77 GluGlyAspTyrPheGlyLeuGluPheProAspHisIleValIleThrValITrpleuAsp 96
Db 358 GAATGTGACTTACTTCGGGATGGAATTCAAAATACTCAGCTCACTGAGATTTGGCTTGAA 417
Qy 97 LeuLeuAspProIleValIlyGlnIleArgArgProIlyHisValValIlyVpheVal 116
Db 418 CTTATGAACCCATCATTTAGGCAATAGCAAGGCGCAAGATGGTGGCTTGCTTGCT 477
Qy 117 ValIlyPhePheProProAspHisIleThrGlnLeuGlnGluIleuLeuThrArgTyrLeuPhe 136
Db 478 GTAAATTTTTCACACTGATCCCTGCTCAGCTCAACAABAATATTACAAGATCTTCTTT 537
Qy 137 AlaLeuGlnValIlyGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAla 156
Db 538 GCCCTTGCACTTGAAGAGACCTGCGAGAGGTTTGACCTGTGTCGACACACGCGC 597
Qy 157 AlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAsp 176
Db 598 GCCCTTTCACGCTCCATCTCTGCACTCGAAATTAAGAAATTACGTGAAGACCTGGAC 657
Qy 177 ArgGluHisLeuAlaIlyAsnIlyTyrIleProGlnGlnAspAlaLeuGluAspIlyIle 196
```

Dh 658 CGAGAGCACCTCAAAAGTAAGATATTTGCTGGCCAGCAGCACCTGCTTGAAGAATA 717
Qy 197 ValGluPheHisHisValIleGlyGlnThrProAlaGluSerAspPheGlnLeu 216
Db 718 CTGAATTCATCAGAAACGCGGCGCAACCTGCTGATCGATTCAGATTCAGGTCCTC 777
Qy 217 GluIleAlaArgArgGluMetGlyGlyIleArgLeuHisProAlaValAspArgGlu 236
Db 778 GAATTCCTGAAAGTTGGAAATGTCAGGCATCAGATTTCATAGCTTTCAGACGGAA 837
Qy 237 GlyThrIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
Db 838 GGAACCAAGATTCAGATCGCAGATTCCTCCACATGGGTACTCTCGGTTCCAGCGCACACC 897
Qy 257 LysIleAsnAlaPheAsnThrAlaValAlaArgLeuSerPheLysValGlyAspGlyPhe 276
Db 898 AAATCAACACTTTCAGATGCTCCAGAGTCGTAACCTTCAAGCTTCAAGAGAAAGATTT 957
Qy 277 LeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
Db 958 CTATTCAAATTCATCCAGAGCTTCATGACCTTACAGACACATTCAGATTTTCTTG 1017
Qy 297 AlaSerArgAspPheCysLysSerPheTyrLysIleCysValGluHisValAspPhe 316
Db 1018 GGTAGTAGATGATGATGTAAGAACTTCTGGAAGATTTGTGGAGTAGTACACCTTTT 1077
Qy 317 ArgLeuPheGluGluProLysProLysProLysProAlaLeuPheSerArgGlySerSer 336
Db 1078 AGACTTTGACCAACTTAAGCAAAAGCAAAAGCCCTTCTTCAACCGGCGCTCTCC 1137
Qy 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyLysHis 356
Db 1138 TTCAGATACAGTGGAAACACTCAGAAACAACTAGTATTTTCAAGACAGTGAATG 1197
Qy 357 LysLeuValGlnPheGluArgLysHisSerLysIleHisSerLysLeu 374
Db 1198 AAGAAGATTCATATGAAAGAGCAGACAGACCAACGCTCCGCTCAGACTGCTGACT 1257
Qy 375 AlaSerGlnProThrGluLeuAsnSerGluValLeuGluGlnSerGlnInsThrSer 394
Db 1258 GCAAGCTTACCAAAACAG-----AGC 1278
Qy 395 LeuThrPheGlyGluGlyAlaGluSerProGlyGlyGlnSerCysArgArg 411
Db 1279 ATCTCATTCCTCCAGAGGATGAGGACTCGCGCTCCCATCTTCAGAGAAAGCTTTTAC 1338
Qy 412 -----GlyLysGluProLysValSerAlaGlyGluProGly----- 423
Db 1339 TCGCTCTCTCCCTCACTGCTGCTCCCTCTGCTGCTGCAAGATTAAAGACAGCAGCAGC 1398
Qy 424 -----SerHisProSerProAlaProArgArgSerProAlaGlyAsnValAsp 441
Db 1399 TCCCTCAACAATCCCGAGTTTCTTACGTCAGAGTCAGCTGACAGAGCGCAGTGA 1458
Qy 442 GlyAlaAlaSerAlaProThrGluGluGluGluValValLysAspArgThrGlnGln 461
Db 1459 GCAAGTGGCTGAGGCCCGAC-----AACCATCG 1488
Qy 462 SerLysProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
Db 1489 GCCCAGCCCTCGGCGCCCGCCGCACTCCAGCTGCTGTCAGGCGCTTTCACAGAGAGTCT 1548
Qy 479 HisLeuSerGluLeuSerValAsnSerGlnGlyGlyValAlaProAlaAsnValThrLeu 498
Db 1549 CAGCCTTCTCTCCAGCCGGAAGAGCCCTGAGTGTAGCCCTGCA-----TTT 1599
Qy 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuLysSerProLeuAsn 518
Db 1600 CAGGTGCTTTGGCCAGCAGCAAGAGGCTCATCCCATCTCGAGCCCTGCTCTCAGT 1659
Qy 519 AspGlnAlaCysProArgThrAspAspGluAspGluGlyArgGlyValArgPheProThr 538
Db 1660 GATGCTGGCGAGCCGGAGTGAAGTCTGAG-----GAGCCAGACACAGAGCGCTGCTGCA 1716

Qy 539 AspLysAlaTyrPheIleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAsp 558
Db 1717 GACAGGCTTACTTCTTATGTCAAAGATTTCTGCTACAGACCAATCTCCAGAT 1776
Qy 559 LeuGluValIleThrSerTyrPheGlnSerThrValSerLysGluAspAlaMetProGlu 578
Db 1777 TTGAAGATTATTCCTGTGGTTCCGACGCGAGTGATGAAGAGCGCATGCTGCGC 1836
Qy 579 AlaLeuLysSerLeuIlePheProAsnPheGluProLeuHisValSerPheHisThrAsnPhe 598
Db 1837 ACTGTATGACGCTGTCTTCTCAACATGCAATCCATCTATATGATTCACAGAGCTTC 1896
Qy 599 LeuLysGluIleGluGlnArgLysAlaLeuTyrGluGlyValArgSerAsnAlaGlnIleArg 618
Db 1897 CTGCGGAGGTGAGCAGAGCTGGCATCTGGGAAGGCGCTCCAAAGCCACAAAA 1956
Qy 619 ---AspTyrGlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLeu 637
Db 1957 GGCAGTCATCAAGATCGGGGACATCTGCTCAGAAACATGCGCCAGTTAAAGAGTTT 2016
Qy 638 AlaHisIleLeuTyrHisHisSerGluAlaLeuGluAlaLeuGluAsnGlyIleLysSer 657
Db 2017 ACCAGCTACTTCCAAAGACATGACGAGTCTTAACAGAACTGGAAGGCTCAACACGC 2076
Qy 658 SerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuPro 677
Db 2077 TGTAAAGATTGAGGACAGTGTACAGAGTTTGACTGCAAGAGCTCTCTCTGCT 2136
Qy 678 LeuAsnThrPheLeuLeuArgProLeuHisValGluLeuMetHisTyrLysGlnValLeuGlu 697
Db 2137 CTCAAACAGTTCCTGCTGAAGCCCATCCAGCGGCTGCTGACATACCGCTGCTGCGC 2196
Qy 698 ArgLeuCysLysHisHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeu 717
Db 2197 CGCTATGCGGCAATTACAGCCCCGGGACACATAGCTACGTCGATGACAGCGCTG 2256
Qy 718 AlaGluIleThrGluMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPhe 737
Db 2257 AAGGCATCAGAGGTCAGACACCACTACAGCAATCTCTCCGGCTGGAGAACTG 2316
Qy 738 GlnLysLeuHisGluLeuLysLysAspLeuIleGlyIleAsnAsnLeuValAlaProGly 757
Db 2317 CAGAACTTAACGAGGCTGACGCGGACCTGTGGCATGAGAACTTCATTCCTGCTGCGC 2376
Qy 758 ArgGluPheIleArgLeuGluLysLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArg 777
Db 2377 AGGGAATTCATCCGTAGAGGCTGCTTCACAGACTCACAAAGAGGCTGCGACGAGG 2436
Qy 778 MetPhePheLeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsn 797
Db 2437 ATGTTTTCCTGCTTCACAGTATGTTGCTGTACAAAGAAAGAGTTCGAGGAGCACAGC 2496
Qy 798 GlnPheLysValHisGlyLeuLeuProLeuTyrGlyMetThrIleGluGluSerGluAsp 817
Db 2497 CACTTCGGATCCGGGGCTCTCTCCCTCCCAAGGCATGCTGTGGAAAGAAATGATTAAC 2556
Qy 818 GluTyrGlyValProHisCysLeuThrLeuAspGlyGlnArgGlnSerIleIleValAla 837
Db 2557 GAGTGTCTGTTCACACGTGTTCACCATCTACGCGGCTCAGAAACAACTGGTGAGCA 2616
Qy 838 AlaSerSerArgSerGluMetGluLysTyrValGluAspIleGlnMetAlaIleAspLeu 857
Db 2617 GCCAGCACTCGCTGAGAAAGAGATGATGCTGACCTGAACCTCCGGATCTCAAGCA 2676
Qy 858 AlaGluLysSerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
Db 2677 GCCAAGATGTCGAGTACACGCGCTGCACTCCAGCGCGACATCTGTGCTACCTGCTCC 2736
Qy 874 ProAspAsnLysSerProAspGluAlaThrAlaAlaAspGlnLysSerGluAspAspLeu 893
Db 2737 CCC-----AGATCCCAACAG-----GATCTCTGAGACGAGGATGACAGATGATGCT 2787

Qy 894 SerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgGlyAsnThrMetValHis 913
Db 2788 CGGGGTGTCGGACCTCCCTGGAAGGCGATGCGACACCGGGCCAAACCAATGAC 2847
Qy 914 ValCysTrpHisArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
Db 2848 GTGGTGTGATCCGGAACACCAAGCGTGTCCAGGCGAAGCCACAGTGCAGCTGTGAGAAC 2907
Qy 934 GlnLeuSerGlyAsnIleuLeuAlaGlyPheIleAsnSerAsnGlyTTPGlnIleuLeuTrp 953
Db 2908 CAGCTTTCAGAGATCTGCTTAAGAAAGTTCAAAAAGTCAATGCTGTGCGCAAACTCTGCG 2967
Qy 954 ValValPheThrAsnPheCysLeuPhePheTrpIleSerSerHisGlnAspAsnHisProLeu 973
Db 2968 GTGCTCTTACCAACTCTGTGTTTCTTCTTACAAAATCTATCGAGATGACTCCACTG 3027
Qy 974 AlaSerLeuProLeuLeuGlyTyrSerLeuThrIleProSerGluSerGluAsnIleGln 993
Db 3028 GCCAGCTCCCGCTGCGGCTACAGCGTACAGCGATCCCGAGGAGCGCATGAC 3087
Qy 994 LysAspTrpValPheIleValPheIleAspPheSerHisValTyrTyrPheAlaGluSer 1013
Db 3088 AAAGACTATGTGTTTCAAGCTCCAGTCAAAATCCACGCTTACTTCCGGCTGAGAGC 3147
Qy 1014 GluTyrThrPheGluArgTrpMetGluValIleArgSerAlaThrSerSerAlaSerArg 1033
Db 3148 AAGTACATTTGAAAGGTGATGAGATGATTCAGGGGGCCAGACAGCTGACCGGAGG 3207
RESULT 11
LOCUS CQ725044 3997 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 10978 from Patent WO2068579.
ACCESSION CQ725044
VERSION CQ725044.1 GI:42285901
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 10978 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source 1..3997
location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
pred. No.: 3,66e-231 Length: 3997
Score: 2936.50 Matches: 580
Percent Similarity: 69.9% Conservative: 161
Best Local Similarity: 54.7% Mismatches: 262
Query Match: 53.8% Indels: 57
DB: 2 Gaps: 14
US-09-555-342b-2 (1-1045) x CQ725044 (1-3997)
Qy 1 MetGlyGlnIleGluGlnArg-----ProThrProGlySerArgLeuGlyAla 16
Db 118 ATGGGGAGATAGAAAGAACTACAGAGTCTGCAAGCTCAGGAGATGCCCTTGAGGCC 177
Qy 17 ProGluAsnSerGlyIleSerThrLeuGlnArgGlyGlnIleProProThrProSer 36
Db 178 CAGACCCCTGTGGAGATGACCTTGCAGCTGGGAGACTCTTGTGCCAAGATGCA 237
Qy 37 GlyLeuValSerIleIleGlnMetLeuAspThrGlnAlaPheGluVal 56

Db 238 GAGAGCACTGCACCTCAAGATTAAGCTGTGCAACAACCAATGAAATTTGACAT 297
Qy 57 ProGlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuVal 76
Db 298 GAGCTAAATGCGATGCGCCAGATTTACTGACACAAGTGTGAAGCGTTTAACTGTGA 357
Qy 77 GluGlyAspTrpPheGlyLeuGluPheProAspHisIleValIleThrValITrpleuAsp 96
Db 358 GAATGTACTACTTGGGAAATGAGTTTCAAAATCTAGTCTTACATGATTTGGCTTGAA 417
Qy 97 LeuLeuLysProIleValIleGlnIleArgArgProLysHisValValIleLysPheVal 116
Db 418 CCAATGAACCCATCATTAAGCAAAATGCAAGGCCAAAGATGTGCTTGGCTTGGCT 477
Qy 117 ValIlePhePheProProAspHisIleThrGlnLeuGlnGluLeuThrArgTyrLeuPhe 136
Db 478 GTAAATTTTTCACCTGATCTGCTGACCTCAAGAAATATATACAGATTAATCTGTT 537
Qy 137 AlaLeuGlnValIleGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAla 156
Db 538 GCTTGCAACTTAAAGAGACCTGCGGAAGAGCGTTTGACCTGTGACACACACAGCG 597
Qy 157 AlaLeuLeuIleSerHisIleValGlnSerGlnIleGlyAspPheAspGluAlaLeuAsp 176
Db 598 GCCCTTCAAGCTCCATCTCTGCGAGTCCGAAATGAGATTTACATGAAGCGCTGAC 657
Qy 177 ArgGluHisLeuAlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIle 196
Db 658 CGAGAGCACTCAAGAGAAAGCAAGTATTTCTGCGCAGACGACACTCCTTGAAGATGA 717
Qy 197 ValGlnPheHisIleAsnHisIleGlyGlnThrProIleGluSerAspPheGlnLeuLeu 216
Db 718 CTAGAAATTCATCAAGAAAGACGTGGGCAACACCTGTGATGCGATTTCCAGTCTC 777
Qy 217 GlnIleAlaArgArgLeuGluMetTyrGlyIleArgLeuHisIleProAlaLysAspArgGlu 236
Db 778 GAATTCCTCGAAGAGTTGGAATGATTCAGCGCATCAGATTTCAATGCTTGCACAGGAA 837
Qy 237 GlyThrIleIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnIlePheThr 256
Db 838 GGAACCAAGATTCAACGCGAGTTTCCCATGATGGTGTACTCGTGTCCAGGCGACACAC 897
Qy 257 LysIleAsnAlaPheAsnThrAlaLysValAlaGlyLeuSerPheLysAspLysArgPhe 276
Db 898 AAATCAACACTTTCACATGCTGTCAGAGTCCGTAAGCTTCAAGAGGAAAGATTT 957
Qy 277 LeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
Db 958 CTTATCAAACTTCATCCAGAGCTTCATGACCTTACCGAGACATTTGAATTTTGTG 1017
Qy 297 AlaSerArgAspPheCysLysSerPheTrpLysIleCysValGlnHisIleAlaPhePhe 316
Db 1018 GTAGTAGATGAATGATTAAGAACTTCTGGAAGATTTGTGAGAGTACACACTTTT 1077
Qy 317 ArgLeuPheGluGluProLysProLysProLysProValLeuPheSerArgLysSerSer 336
Db 1078 AGACTTTTGACCAACCTTAAGCCAAAGCAAAAGCCGTCTTTCACCGGGCTCTCC 1137
Qy 337 PheArgPheSerGlyArgThrGlnIleGlnValIleuAspTrpValIleGlyGluGlyHis 356
Db 1138 TTCAATACAGTGGAGAACTCAGAAACAACTAGTATTTTTCAAAGACAGTGAATG 1197
Qy 357 LysLeuValGlnPheGluArgLysHisSerLysIleHis--SerIleArgSerLeu-- 374
Db 1198 AAGAGAAATTCATATGAAAGAGCAACAGCAACCAACGATCGCTGAGCTGACT 1257
Qy 375 AlaSerGlnProThrGluLeuAsnSerGluValLeuGluGlnSerGlnInsThrSer 394
Db 1258 GCAAGCTTACCAAAACG-----AGC 1278
Qy 395 LeuThrPheGlyGluGlyAlaGluSerProGlyGlyGlnSerCysArgArg----- 411

Db 1279 ATCTCATCCCGAGGATTGAGGACTCTGCTCCCATCTTCAGCGAATGCCCTTTAC 1338
 QY 412 -----GlyylserIuProlyValserAlaGlyIuIuProGly----- 423
 Db 1339 TCGCTCTCTCCCTCCACTGTGTCCTCTGCGCTGCGAAGTTTAAAGACAGCAGCAGC 1398
 QY 424 -----SerHlaProSerProAlaProArgArgSerProAlaGlyValnlySerAlaAsp 441
 Db 1399 TCCCTCACAAGATCCCAAGCTTCTCTACGTCAAGAGTCACCTCAGAGAGCCCAAGTGA 1458
 QY 442 GYVAlAlaAserAlaProThrGluGluGluGluValVallyAspArgThrGln 461
 Db 1459 GCATGGCTTGAGAGCCCGAC-----ACACATCG 1488
 QY 462 SerLyProGlnPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
 Db 1489 GCCACGCTCGAGGCCCCCGCAGCTTCACGCTGTGTCAGAGCCTTTCAGAAAGCTCT 1548
 QY 479 HlaLeuSerGluLeuSerValaanserGlnGlyValAlaProAlaAsnValThrLeu 498
 Db 1549 CAGCTTCTCCCTCCAGCCGGAAGAGCCCTGAGTCTAGCCCTGCA-----TTT 1599
 QY 499 SerProAsnLeuSerProAspThrLyGlnAlaSerProLeuIleSerProLeuAsn 518
 Db 1600 CAGCTCCTTGGGCGCCAGCTGAACAGGCTCATCCCTGAGCCCTGTCTCAGT 1659
 QY 519 AspGlnAlaCybProArgThrAspArgGluAspGluGlyValArgArgLyAspArgProThr 538
 Db 1660 GATGCTGGCGGAGCCGAGATGACTGCGAG--GAGCCAGACACAAAGCCGTGCTGCA 1716
 QY 539 AspLyAlaLyThrPheIleAlaLyGluValaSerThrThrGluArgThrLyLeuLyAsp 558
 Db 1717 GAGAGGCTCTACTTCACTAGCAAGAGATCTGCGTACAGAACATATCTCAAGAT 1776
 QY 559 LeuGluValIleThrSerThrPheGlnSerThrValSerLyGluAspAlaMetProGlu 578
 Db 1777 TTGAAGATTATTCCTGTGTGTTCCGACGCGAGTGTGAAGAGGACGCTCATGCTGCG 1836
 QY 579 AlaLeuLySerLeuIlePheProAsnPheGluProLeuHlaSerPheHlaSerAsn 598
 Db 1837 ACTGTATGACGGTGTCTTCTCCAACTGATCCCATATGAGTTCCACAGAGCTTC 1896
 QY 599 LeuLyGluIleGluGlnArgLeuAlaLeuTrpGluGlyValArgSerAlaGlnIleArg 618
 Db 1897 CTGCGGAGGTGAGCAGAGGCTGCGCACTGCGAAGGCCCTCCAAAGCCACACAAA 1956
 QY 619 ---AspTyTrpGlnArgIleGlyAspValMetLeuLyAsnIleGlnGlyMetLyHlaLeu 637
 Db 1957 GGCAGTCACTCAAGAAATCGGGGACATCTGCTCAGGAACATGCGCCAGTTAAAGAGTTT 2016
 QY 638 AlaAlaHlaLeuTrpLyHlaSerGluAlaLeuGluAlaLeuGluAlaGlyIleLySer 657
 Db 2017 ACCAGCTACTTCCAAATACATGACGAGGTCTTAACAGAACTGGAAGAGCTACCAAGC 2076
 QY 658 SerArgArgLeuGluAsnPheCybArgAspPheGluLeuGlnLyValCybTyTrpLeuPro 677
 Db 2077 TGTAAAGATTGGAGGACAGTGAAGAGATTGAGCTGGAAGAGGTCTGCTTGGCT 2136
 QY 678 LeuAsnThrPheLeuLeuArgProLeuHlaValGluLeuMetHlaTyTrpLyGlnValLeuGlu 697
 Db 2137 CTCMAACAGTTCTCTGTGAAGCCCATCCAGGCTGTGCTGCTGCTGCTGCTGCTGCTG 2196
 QY 698 ArgLeuCybLyHlaSerProSerArgHlaAspPheArgArgCybArgAlaAlaLeu 717
 Db 2197 CGCTGTGCGACATTACAGCCCGGCGGACCATGATCACTGACATGACGCGCTG 2256
 QY 718 AlaGluIleThrGluMetValaAlaGlnLeuHlaGlyThrMetLySerGluAsnPhe 737
 Db 2257 AAAGCATACAGAGGTGACACACACATCAAGACATCTATCCGCGTGGAGAACTCG 2316
 QY 738 GlnLyLeuHlaIleGluLeuValyValAspLeuIleGlyIleAspAlaLeuValaProGly 757
 Db 2317 CAGAACTAAACGAGCTGCGAGCGGACCTGTGTGGGATAGAAACCTCATTTGCTCTGGC 2376

QY 758 ArgGluPheIleArgLeuGlySerLeuSerLyLeuSerGlyLyGlyLeuGlnIleArg 777
 Db 2377 AGGAGATTATCCGTAGAGGCTGCTTCAACAGCTCAACAAAGAGGCTGCGACGAGG 2436
 QY 778 MetPhePheLeuPheAsnAspValLeuLeuTyThrSerArgGlyLeuThrAlaSerAsn 797
 Db 2437 ATGTTTTCTGTGTTCATATATGTGTGTACACAAAGAGAGTGTAGGAGGACAGC 2496
 QY 798 GlnPheLyValHlaIleGlyLeuProLeuTyTrpGlyMetThrIleGluGluSerGluAsp 817
 Db 2497 CACTTCGGATCCGGGCTCTCTCCCTCAAGCATCTGTGTGGAAGAAATGTATAC 2556
 QY 818 GluTrpGlyValProHlaCybLeuThrLeuAspGlyGlnAlaGlnSerIleIleValAla 837
 Db 2557 GAGTGTCTGTTCACACACTGTTTCACTACATCTACGCGGCTCAGAAAACATCTGTGGCA 2616
 QY 838 AlaSerSerArgSerGluMetGluTyTrpValGluAspIleGlnMetAlaIleAspLeu 857
 Db 2617 GCCAGCACTCGGCTGGAGAAAGAAAGTGAATGTGATCTGATCTGATCCGCGATCCAAGCA 2676
 QY 858 AlaGluLySerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
 Db 2677 GCCAAGATGGCGGTGACACGCGCTGCACTGCCAGCGCGCACTGTGTGCATCGTCCC 2736
 QY 874 ProAspAsnLySerProAspGluAlaThrAlaAlaAspGlnGluSerGluAspAspLeu 893
 Db 2737 CCC-----AGATCCCCCAACAG--GATCTCTGAGACAGAGTCAAGATGTGCT 2787
 QY 894 SerAlaSerArgThrSerLeuGluArgGlnAlaProHlaArgGlyValanthMetValHla 913
 Db 2788 CGGAGTGTCCGAGACTCCTTGAGGGGAGATGGCAGACACCGGCGCAACACACATATGCAC 2847
 QY 914 ValCybTrpHlaArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
 Db 2848 GTGTGTGTGTACCGGAACACAGCTGTCCAGGCGACACACAGTGCACCTGTGAGAAC 2907
 QY 934 GlnLeuSerGlyAsnLeuLeuArgLyPheLyAsnSerAsnGlyTyTrpGlnLyLeuTrp 953
 Db 2908 CAGCTTCAAGATATCTGTGAAGAAATTCABAACAGTCAAGCTGCGAGAAAGCTCTGG 2967
 QY 954 ValAlaPheThrAsnPheCybLeuPhePheTyTrpSerHlaGlnAspAsnHlaSerProLeu 973
 Db 2968 GTGCTTTTACCAACTTCTGTTGTCTTCAAAAATCATCAAGATGATGATCAACCATG 3027
 QY 974 AlaSerLeuProLeuGlnGlyTySerLeuThrIleProSerGluSerGluSerAlaArg 993
 Db 3028 GCACGCTTCCGCTGTGGCTTACAGCGTGAACATCCCAAGAGGCGCATGGCATACAC 3087
 QY 994 LyAspTyValPheLyLeuHlaSerHlaSerHlaValTyTrpPheArgAlaGluSer 1013
 Db 3088 AAAGACTATGTTTCAAGCTCAAGTTCAAATCCCAAGTCTACTTCTCCGGGTGAGAGC 3147
 QY 1014 GluTyTrpThrPheGluArgTrpMetGluValIleArgSerAlaThrSerSerAlaArg 1033
 Db 3148 AAGTACACATTTGAAAGGTGATGAGGTGATTCACAGGCGGCAGACTCAGCGGAGAG 3207
 RESULT 12
 BC009153 3719 bp mRNA linear ROD 29-JUN-2004
 LOCUS Mus musculus FERM, Rhogef and plectristin domain protein 2, mRNA
 DEFINITION (CDNA clone MGC:6304 IMAGE:2655209), complete cds.
 ACCESSION BC009153
 VERSION BC009153.1 GI:14318718
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Mus.
 1 (bases 1 to 3719)
 Struhsberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marinova, K., Farmer, A.A., Rubin, G.M., Hong, L., Slaughter, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Tothiyuk, S., Carninci, P., Pringle, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mulle, J.S., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Feeney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E., Schmechel, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
2 (bases 1 to 3719)
Strauberg, R.
Direct Submission
Submitted (05-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H., Kowles, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LML at: <http://image.llnl.gov>
Series: IRAC Place: 5 Row: m Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21704021.
location/Qualifiers
1. 3719
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="MGC:6304 IMAGE:2655209"
/tissue_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month old mouse, Taken by biopsy."
/clone_lib="NCI_CGAP_Mam2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 3719
/gene="Farp2"
/note="synonym: Flr"
/db_xref="GeneID:227377"
/db_xref="MGI:2385126"
99. 3296
/gene="Farp2"
/codon_start=1
/product="FERM, Rhogp and plectectrin domain protein 2"
/protein_id="AAH09153.1"
/db_xref="GI:14318719"
/db_xref="GeneID:227377"
/db_xref="MGI:2385126"
/translation="MGEIEGTVALPTSGTRLGQGTAAIGVSTLEPEQSLSPRQEKHM

ORIGIN

Alignment Scores:
Pred. No.: 3,55e-230 Length: 3719
Score: 2924.00 Matches: 563
Percent Similarity: 70.2% Conservative: 179
Best Local Similarity: 53.3% Mismatches: 267
Query Match: 53.5% Indels: 48
DB: 6 Gaps: 9
US-09-555-342b-2 (1-1045) x BC009153 (1-3719)
QY 1 Merglvgluilegu-----GlnargProthProglyserArgleuGlyala 16
99 ATGGAGAGATTAAGAGACATACAGAGCTTACCACTTCAGAGACCTTGAGGCTGC 158
QY 17 ProgluamserglyleaserthleugluargglylnlysbProprothProser 36
159 CAGACCGCATTTGAGAGATACCTTGAGCCAGAGAGAGCGCTGACCCAGAGATG 218
QY 37 Glylyleuvalserllelyslleglmetleuaspaspthrglngluualpneglua 56
219 GAGAGACACATGCGCATCAGAGTGAAGCTGCGACGACCTGATGATTATGACATC 278
QY 57 ProgluarglalarProglylsvalleuaspalaValysasnhsleaualeuVal 76
279 GAGCCGAGGTACCGGACGCTTACGACCTGACCAAGAGAGAGAGAGAGAGAGAG 338
QY 77 GluglyAspTyrrheglyleugluupheProaspshlelysllethValTrrpleuasp 96
339 GAATGTGACCTACTTGCTGAGATTCAAGATGTCAGATGCTGAGATGAGATGAGAT 398
QY 97 LeuleuylsbProilleVallysglnlleargatProlyshlsValVallysbpheVal 116
399 CCAATGAAACCACTATCATTAGCAAGTACGAAACCAAAATATCGGTGCTTGGCA 458
QY 117 VallysbpheheProaspshstnrglnleuglngluileuThrArgTyrrleuPhe 136
459 GTAAATTTTTCCTGCTGACCTGCTGACCTGACCTGACCAAGAGAGTACACAGAGT 518
QY 137 AlaLeuglnVallysglnaspLeuAaglnGlyArgleuThrCybsasnAspThrsEra 156
519 GCCTTGCAACTCAAGAGAGACCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 578
QY 157 AlaLeuLeuileserthslleValGlnserGluileGlyAspPheaspGlualeuasp 176
579 GCCCTTCTCATATCCACCTTCGACGTCGGAATGGAATATGATGATGAAACCTGGAT 638
QY 177 ArgGlnHisleuAlalyAsnlystYrrleProGlnGlnAspAlaleuGlnAspLy 196
639 CGAGAACCTCAAG 658
QY 197 ValGlnPheHisAsnHisleGlyGlnThrProAlaGlnAspAspPheGlnleu 216
699 CTGAGACTTCATCAGAGGACAGCGGCGAGAGACTCGAGAGAGTCAAGATTCAGAGT 758

proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntrophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"

/db_xref="CD:smart0023"

ORIGIN

Alignment Scores:

Pred. No.:	6,45e-164	Length:	2632
Score:	2117.00	Matches:	400
Percent Similarity:	97.6%	Conservative:	10
Beet Local Similarity:	95.2%	Mismatches:	10
Query Match:	38.8%	Gaps:	0
DB:	6		0

US-09-555-342b-2 (1-1045) x BC004009 (1-2632)

```

QY 626 ValMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaHisLeuTrpLysHisSer 645
DB 1 GTGATGCTGAAAAAATTCAGGGCAGTAAACACCTGGCGGCCCACTGTGGAAGCACAGT 60
QY 646 GlnAlaLeuGlnAlaLeuGlnAlaGlnGlyIleLysSerSerArgArgLeuGlnAlaAspPheCys 665
DB 61 GAGGCGCTGAGGCGCTGGAGACCTCCATCAAGGCGCTCGGCGGCGCTGGAGACATTTCTGC 120
QY 666 ArgAspPheGlnLeuGlnLysValCysTyrLeuProLeuAsnThrPheLeuLeuArgPro 685
DB 121 CGAGACTTCGAGCTGCAAGAGGTGTGTACTCTCGCTCAACCTTCCCTCGCGCCCG 180
QY 686 LeuHisArgLeuMetHisTyrLysGlnValLeuGlnLysGlnCysValHisAspPro 705
DB 181 CTGCACCGGCTCATGCACTAATGATGCTCTGAGAGGCTGTGCAAGCACCAACCA 240
QY 706 SerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGlnIleThrGlnMetValAla 725
DB 241 AACCAACCGGCACTTCAGGAGCTCAGAGCTGCCCTGGCGGAGATCAAGAGATGTGGCC 300
QY 726 GlnLeuHisGlyThrMetIleLysMetGlnAspPheGlnLysLeuHisGlnLeuLysVal 745
DB 301 CACCTGCACGCGACCATGATCAAGATGGAACATTCAGAAAGCTGCATGAGCTCAAGAAA 360
QY 746 ArgLeuIleGlyIleAspAsnLeuValValProGlyArgGlnPheIleArgLeuGlnSer 765
DB 361 GATCTGATCGGCAATGACAACTTGTGATCCAGAGAGGAGATTCACTCCGCTGGGAGC 420
QY 766 LeuSerLysLeuSerGlyLysGlyLeuGlnGlnArgMetPhePheLeuPheAsnAspVal 785
DB 421 CTCAGCAAGCTCTCCGGGAGGGGCTTCAGCAGGCAATGTTCTTCCTGTTCAAGATGTC 480
QY 786 LeuLeuTyrThrSerArgIleLeuThrAlaSerAsnGlnPheLysValHisGlyGlnLeu 805
DB 481 TTCTCTGATATCCAGCCCGGAGCTGACAGCATCTAATCAAGTTAAAGTCCACGACAGCTC 540
QY 806 ProLeuTyrGlnMetThrIleGlnLysSerGlnAspGlnTrpGlyValProHisCysLeu 825
DB 541 CCACCTCATGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 826 ThrLeuArgGlnArgGlnSerIleIleValAlaHisLeuSerSerArgSerGlnMetGln 845
DB 601 ACCCTGAGGGGCGAGCGGCAATCATGTGCTCCAGTTCCCGGTCAAGAGATGAA 660
QY 846 LysTrpValGlnAspIleGlnMetAlaIleAspLeuAlaGlnLysSerSerSerProAla 865
DB 661 AAATGATGATGAGACATCCAGATGGCATTGACCTGGCAGAGAAAGCAATGACCCACC 720
QY 866 ProGlnPheLeuLysSerProProAspAsnLysSerProAspGlnAlaThrAlaAla 885
DB 721 CCGAGCTCTTGCCAGACAGCCCTCCGACAAAGTCCCGAGATGAGAGCAAGCGGCT 780

```

```

QY 886 AspGlnLysSerGlnAspAspLeuSerAlaSerArgThrSerLeuGlnAlaPro 905
DB 781 GACCAAGAGATCAGAGAGATACCTGAGCGGCTTCGAGATCCTGGAGCGCCAGGCCCCCT 840
QY 906 HisArgGlyAsnThrMetValHisValCysTrpHisArgAsnThrSerValSerMetVal 925
DB 841 CACCGTGGGCAACAGATGTGACGTGTGCGATCGGAGCACAGTGTCTCCAGATG 900
QY 926 AspPheSerIleAlaValGlnAsnGlnLeuSerGlyAsnLeuLysArgPheLysAsn 945
DB 901 GACTTCAGCATTCGCTGTGAGAAACAGCTCTCGGGAATCTCTAGAAATTCAGAAAC 960
QY 946 SerAsnGlyTyrGlnLysLeuTrpValValPheThrAsnPheCysLeuPhePheTyrLys 965
DB 961 ACCAATGCTGGCAGAGAGCTGTGGTGTGATTCACCAACTTGTGCTGTTCTTCAAAA 1020
QY 966 SerHisGlnAspAsnHisAspLeuLysSerLeuProLeuGlnGlyTyrSerLeuThrIle 985
DB 1021 TCACACAGAGACAGTACCCCTTGCAGGCTGCTGTGAGGCTACTCCCTCACCATC 1080
QY 986 ProSerGlnSerGlnAsnIleGlnLysAspTyrValPheLysLeuHisAspPheLysSerHis 1005
DB 1081 CTTCTGAGTCTGAGAAACATCCACAAGACATATGTTCACACTGCATTTCAAGTCCAC 1140
QY 1006 ValTyrTyrPheArgAlaGlnSerGlnTyrThrPheGlnArgTyrMetGlnValIleArg 1025
DB 1141 GTGTACTACTTCAGGCGCTGAAGCAATACATTTGAAGGTGGATGGAATGATCCGA 1200
QY 1026 SerAlaThrSerSerLysSerArgProHisValLeuSerHisGlyGlnSerLeuValTyr 1045
DB 1201 AGTGCCACAGCTCAGCCCTTGCGGCCACATCTTGATGACAAAGAAATTCACCTGTAC 1260

RESULT 14
BT024186
LOCUS BT024186 4048 bp mRNA linear INV 28-JAN-2006
DEFINITION Drosophila melanogaster SD10794 full insert cDNA.
ACCESSION BT024186
VERSION BT024186.1 GI:85857423
KEYWORDS FLI CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 4048)
REFERENCE
  Stapleton,M., Carlson,J., Chavez,C., Frise,E., George,R.,
  Pacleb,J., Park,S., Wan,K., Yu,C. and Ceiniker,S.
  Direct Submission
  Submitted (28-JAN-2006) Berkeley Drosophila Genome Project,
  Lawrence Berkeley National Laboratory, One Cyclotron Road,
  Berkeley, CA 94720, USA
  Sequence submitted by:
  Berkeley Drosophila Genome Project
  Lawrence Berkeley National Laboratory
  Berkeley, CA 94720
  This clone was sequenced as part of a high-throughput process to
  sequence clones from the Drosophila Gene Collection. The sequence
  has been subjected to integrity checks for sequence accuracy,
  presence of a polyA tail and contiguity within 100 kb in the
  genome. Thus we believe the sequence to reflect accurately this
  particular cDNA clone. However, there are artifacts associated with
  the generation of cDNA clones that may have not been detected in
  our initial analyses such as internal priming, priming from
  contaminating genomic DNA, retained introns due to reverse
  transcription of unspliced precursor RNAs, and reverse
  transcriptase errors that result in single base changes. For
  further information about this sequence, including its location and
  relationship to other sequences, please visit our web site
  (http://www.fruitfly.org) or send email to cdna@fruitfly.org.
  Location/Qualifiers
    1..4048
    /organism="Drosophila melanogaster"
    /mol_type="mRNA"
  FEATURES
    source

```

CDS

```

/db xref="taxon:7227"
2.3505
/note="longest ORF"
/codon_start=1
/product="SD10794p"
/protein_id="ABC6248.1"
/db xref="GI:85857424"
/translation="MSLADMGTAERSAGGEGGRHYDLAGTSGSGGHPIVGGIPIGGRM
NLBADIYGLBYGEVSTHTRKWLDELEPKMKRQVSLIDPLVRFCTFYIPDPQQLS
EYRYFLCLOIKDLATGSLQCDNDNTAALASYVOASCDFPEVDYDHTYLSYF
VPQDAMORKIMENHKNHVAQSPAEADNLLETARCEJLYGMMHPAKQVEGPLN
AVAMGITVFONIRINTFSMAKIRKISFKRKRLVLAHEGYVDQDTYAFEEGN
ECKPMKCVENGEFRTCAVONTPRKRTIARGSSFRYSQKOKOITFEVENYK
RONRPOSFPROGPIASRSOSHTYNSISANPLIPITAMPDYNOCSDSMTPEL
TKKADLTDRRDNPIGHRMSQVTAQVETIQTNYADSPTSOEBARCSAERCH
HSAVAMDKLNSNRSLSPQSPQSWTSPSHSHHQRPDPARVHPGDNTGTGYGINA
MSLDRGEITTPPRYDLITLGSKSSLSSEAGTYDVAIEIOHAROBELIAGVANA
SHQNGNGNGHTLSTOHDIEAEVKRKPTEPSYFLAKELMTERYKKDLVLAHTT
PROVLSGDVEOLQPLFELLDSLQHNLLFLROIEMHVMQEGRGHEARIGDVMWK
HMAALPYDERVYVTHLIDILHCDNMVREGDERPROVYKEPQKCYPIBELLKPLN
RLAHYQILIRLCDYGEENIDYADAAVHHLLVSTKGRISQPLDSANVELCELDK
DINEOLVOPHRLIRQGCCLKSKRGIAQRMPELFSDDLKYGSKSPIDRFRIGHV
PVSLLTENAEHNTFSIFGQCAITVAGTAETKTLAELSKAADIDKRPMPMOQL
LTLKNGSSSEEGDLFGLSNGNNSLNSSNGGPLTTOOKLOLQOONRTPSR
SNRAHYCMHKGATVGLGDHLIAEHOLSGYLLRKFNKSMOKLWVPTSPCLYFYK
SYODRPLASLPLIGTYVPRPGHODAVOKEPVFKLSFKMHVYFFRAESATYNNMLEV
LRSITOTIDFKNVSHAVLGN"

```

ORIGIN

Alignment Scores:

```

Pred. No.:      4 24e-144      Length:      4048
Score:          1879.50      Matches:      437
Percent Simlarity: 54.4%      Conservative: 197
Beet Local Simlarity: 37.5%      Mismatch: 363
Query Match:      34.4%      Indels: 169
DB:              13          Gaps: 28

```

US-09-555-342B-2 (1-1045) x BT024186 (1-4048)

```

QY 8 ProThrProglYserArgLeuGlyAlaProgluAenSerGlyIleSerThrLeuGluArg 27
Db 101 CCAATGGCGCGCGGTCTGCCGAGCGCGCATGACGACCTCCACACACTCTTCGGG 160
QY 28 GlyGlnYserProProThrPro-----SerGlyLeuValSerIleGlyIle 44
Db 161 GTGGATGGACGCGCTCCACCCGACATCGCGCGCGCAAAAAGCTGACGTTCGCATC 220
QY 45 GlnMetLeuAerAerThrGlnGluAlaPheGluValAProgluArgAlaProglYLeuVal 64
Db 221 CAGATGCTGACCACTGCATGCATGCTTCACATGACGCTAAAGCACTGCGCGCGGTG 280
QY 65 LeuLeuAerAlaValCyAaenHISLeuAenLeuValGluGlyAerTYrPheGlyLeuGlu 84
Db 281 CTGTTTGAACAGGATACCGCTCAGCTGATTTACTGAGGCGGACCTACTTCGCTTGAG 340
QY 85 PheProAerP-----HISLysLysIleThrValIlePheAerLeuLeuYProIle 101
Db 341 TACCAAGGAAGTCTCCACACATACCAA-----TACTGGCTGATCTCGAGAAAGCCATG 394
QY 102 ValLysGlnIleAerArgProLys---HISValValValLysPheValValLysPhePhe 120
Db 395 AACCGGCAAGGTGGCTCTCTTATGATTCAGATGCTGCGCTTCTGCATCAAGTTTAT 454
QY 121 ProProAerPHeThrGlnLeuGlnGluGlyLeuThrArgTYrLeuPheAlaLeuGlnVal 140
Db 455 AACCCGATCCAGCGCATTTGAAAGAGAGATGACCAAGATATTGTTCTGCTTCAATC 514
QY 141 LysGlnAerLeuAlaGlnGlyArgLeuThrCyAaenAerThrSerAlaAlaLeuLeuIle 160
Db 515 AAACGAAACCTGGCCACAGGCACTCTCAGTCAACATTAACACGCGGCTTATATGCA 574
QY 161 SerHISLysValGlnSerGluIleGlyAerPhe-----AerGluAlaLeuAerArgGlu 178

```

```

Db 575 AGCTACATTGTACAGGGGTATGCGGGACTTCTCCGAGAGATATCCGACCATACT 634
QY 179 HISLeuAlaLysAnlysTYrIleProGlnGlnAerAla---LeuGluAerLysIleVal 197
Db 635 TATCTGCTCCGTATCGCTTGTGTCACACAGGACCCACATGCGCAAAAAGATTATG 694
QY 198 GlnPheHISLysAnHISLISLysIleGlnIleThrProAlaGlnYserAerPheGlnLeuLeuGlu 217
Db 695 GAGAACCAAAAAGCATGTGTGTCATATCCCGGAGGCGGACCTTAACCTTTGGAG 754
QY 218 ILeuAerArgLeuGlnMetTYrGlyIleAerGlnHISProAlaLysAerArgGluGly 237
Db 755 ACGGCTCGAGAGGTGTAGACGTATGCGATGAATAATGATCCGGGAAGATGTGAAAGG 814
QY 238 ThrLysIleAenLeuAlaValAlaSerThrGlyIleLeuValPheGlnGlyPheThrLys 257
Db 815 GTGCGGCTTAATGTGGCTGTGGCCCATGCGGATCAAGTCTTCCAGAACATCACCGG 874
QY 258 ILeuAerAlaPheAerThrAlaLysValArgLysLeuSerPheLysArgLysPheLeu 277
Db 875 ATCAACACTTCTCGTGGGCTTAAGATACGAAATTTCTTCAAGCGCAAGCATTCCTG 934
QY 278 ILeuLysLeuAerProAerAlaAenSerAlaTYrGlnAerThrLeuGlnPheLeuMetAla 297
Db 935 GTCAACTGCATCCAGAGGATATGATATTAACAAGATACCGTGCAGTCTTCTTGAG 994
QY 298 SerArgAerPheCyAlysSerPheTrpLysIleCyAValGlnHISLISLysAlaPhePheArg 317
Db 995 GGTGCGCAACGATGCAAAAACCTTCGAAAAAATGCGTCGAAAAATCCGATTCCTCGA 1054
QY 318 LeuPheGlu---GluProLysProLysProLysProValLeuPheSerArgLysSer 336
Db 1055 TGCATGCGCTTACAAAATATACGCCACGCGCAAACTCGCTTCTCGCGGAGTGTCA 1114
QY 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAerTYrValLysGluGlyLys 356
Db 1115 TTCGCTATAGCGGAAAAAACCGAAGCAATATGCAAGTTCGTCGCAAAATTATGTCG 1174
QY 357 LysLysValGlnPheGluArgLysHISerLysIleHISerIleArgSerLeuAlaSer 376
Db 1175 AACGCGCAAACTTCCAAAGCTC-----CAGTATTCGCGCA 1213
QY 377 GlnProThrGlnLeuAenSerGlnValLeuGlnGlnSerGlnGlnIleThrSerIleThr 396
Db 1214 GGGCCTGAATGCTGTAGC-----GAAGTCAATCGCATACGATATGTA 1260
QY 397 PheGlyGluGlyAlaGlnYserProGly-----GlyGln 407
Db 1261 TTCAGATATTCAAGCAATCCCTTCTCCAAATTGACATCGCGCATGCGACATGCA 1320
QY 408 SerCyAerArgGlyLysGluProLysValSerAlaGlyGluProGlySerHISProSer 427
Db 1321 TCAATGACGACGATCGATGACCCC-----TTCCTGACAGAAAGGACGATACCTT 1374
QY 428 ProAlaProArgArg---SerProAlaGlnYenLysGlnAla----- 440
Db 1375 GATTCGCGCAGCGACATCTATTCGACCATGCGCTTCTCAAGTCAACAGCACACCCAGGT 1434
QY 441 -----AspGlyAlaAlaSerAlaProThrGlnGluGlu----- 452
Db 1435 GAGATCTATCAACAGAAAGATATGACAGCGGACCTGCCCACTCCACAGAGAAAGCGGA 1494
QY 453 -----GluValValLysAerArgThrGln 460
Db 1495 ATGCTCGCGCGCGGCGGAGCAACGCAACACATCGCGCATGTATGACAAACTGAA 1554
QY 460 ngInserLys-----ProGlnProProGlnProSerThrGlySerLeuThrGlyse 477
Db 1555 TTGGAACCGCTCACTTTCGCCCAAGGACCGCATGTGTGACT-----TC 1599
QY 477 rProHISLysSerGluLeuSerValAenSerGlnGlyValAlaProAlaAenVal-- 496

```

Db 1600 GCCC-----AGTCACAGACAGCCACCAACGCGGCTCCGATCCGGCTGCTGTGA 1653
 Qy 497 -----ThrluserProApnLeuSerProApn 506
 Db 1654 TCCAGCGCATCACAAATACAGATGGTATATGGAATAAATGATGCTGTGAACAG 1713
 Qy 506 rlyb-----GlnlaserProleu1 513
 Db 1714 GCGCGGTGAATACCACTCCACCAACGCGGTACGACTCCTTAGATCGGATAGTC 1773
 Qy 513 eSerProleuLeuApnGlnlase 521
 Db 1774 TTCGTACATATCCCGATCTGAGGACAGGAACATATGACGTATATACAGGCAAAATCCAGCA 1833
 Qy 521 ----- 521
 Db 1834 CGCTAAGACAGAGAGCTAGCCACAGGCGTTCGACCGCTCGACCAACGAAATGGAAACGG 1893
 Qy 522 -----CysProArGThrApnApnGluApnGluGlyArGArGly 534
 Db 1894 AAACGGAAATGGGCATACCTCTGACACACACACATATGAGGAGAAATGAAAAAGG 1953
 Qy 534 sArGpHeProThrApnGlyAlaTyRPhelAlaTyRGlValSerThrThrlGluArgTh 554
 Db 1954 GAAATGGCCACCGACCCGACTATTTTGGCCAAAGAGCTGTGATGACGGAAACGTAC 2013
 Qy 554 rTyRLeuValApnLeuGluValIleThrSerTrpPheGlnSerThrValSerIyGluVal 574
 Db 2014 GTACAAAGAGATCTGATGTGCTGAACACACACCTCCGACAGCTGTGATGCTCGAGGA 2073
 Qy 574 PalmetProGluValAlaLeuValSerLeuIlePheProApnPheGluProleuValIyRph 594
 Db 2074 TGGG-----GAGCAGGTGACGACCGCGT-----TTCAGAGCTGTGACTCCCTGGCCACGA 2124
 Qy 594 eHlThrApnPheLeuValIleGluGlnArgLeuAlaLeuTrpGluGlyArgSerAla 614
 Db 2125 CCACAATCTCTCTCTGCGACATCGAACCCGACCTGTGCACTGGAGGAAAGAGAGG 2184
 Qy 614 nAlaGlnIleArgApnTyRGlNArgIleGlyArgValMetLeuValApnIleGlnGlyMe 634
 Db 2185 C-----CACGAAGCCACCGCATCGAGACGTCATATGATAGACACATGGCCGCTCT 2235
 Qy 634 rLybHlAlaLeuAlaAlaHlAlaLeuTrpValIleSerGluValAlaLeuGluValApnI 654
 Db 2236 GCCCATCTACGATGAGTACGTGACAGACGACCTGACATCTGACATGTATGAACACAT 2295
 Qy 654 yIleIySerSerArgArGLeuGluValApnPheCybArGApnPheGluGlnIyValCy 674
 Db 2296 GTACGAAGCGCATGAACGCTTCGTCAGGTGTACAGAGCTTTGAGACGAAAAAGGTTTG 2355
 Qy 674 sTyRLeuProLeuApnThrPheLeuValArgProLeuHlArgLeuMetHlTyRlySGl 694
 Db 2356 CTATCTACCATCGCGCACTTCTACCTGAAGCCCCCTTCACCGGCTGTGACTCAACT 2415
 Qy 694 nValLeuGluArgLeuCybAlaIyRphProSerHlAlaApnPheArgArpCybAr 714
 Db 2416 GATCTTAGAGCGGCTGTGACACTATAGGGAGAGACATATGCATATCCGATGTAT 2475
 Qy 714 gAlaAlaAlaValAlaGluIleThrGluMetValAlaGlnLeuHlGlyThrMetIleIySMe 734
 Db 2476 GGCCTGTGACACACTGCTCGTTCGACGACCAACGAGTATAGTCCGAGCTCCGACTC 2535
 Qy 734 rGluApnPheGlnIyLeuHlArgIleuValIyValApnLeuIleGlyIleApnLeuVal 754
 Db 2536 TGCAAACTTGTGAGAGCTGTGATACGTCAACCGGAC--ATCAACTTCGACAGCTTGCT 2592
 Qy 754 lValProGlyArgGluPheIleArgLeuGlySerLeuSerIyLeuSerGlyIySGlyLe 774
 Db 2593 TCAAGCCCATCGCGGCTCATCCGCCAGGATGCTTCGAAAGACTCCAAAGCGGCTCT 2652
 Qy 774 uGlnGlnArgMetPhePheLeuPheApnValLeuLeuTyThrSerArgGlyLeuTh 794
 Db 2653 GCAGCAGAGATGTCTTCTTGTGTTCTCCAGCTGCTCTACGGTTCCGAAGTCC--CC 2709

Qy 794 rAlaSerAenGlnPheIyValHlGlyGlnLeuProLeuTyRGlYMetThrlIleGluI 814
 Db 2710 GCTGACACAGACTTCGCCATTTTAGCCACAGTCTGTCGCTTACTTATCTACCGAGAA 2769
 Qy 814 uSerGluApnGluTyRGlValProHlCybLeuThrlAlaArgIleSerI 834
 Db 2770 CGCCGAGCAC-----AACACTTCCTCATCTTCGCGGAGCAGTGTCCAT 2814
 Qy 834 eIleValAlaAlaSerSerArgSerGluMetGluTyRTrpValGluApnIle--GlnIe 853
 Db 2815 CACAGTAGTGGCGGACACCGCTGAAGAACCTCTGCTGGCCGAGCTTCACAGGC 2874
 Qy 853 rAlaIleApnLeuAlaGluIyS----- 860
 Db 2875 GGCCTGGATATTAACCAACCGCCCAATGCAACTCCAGCTAACGACGCTCAAGAA 2934
 Qy 861 -----SerSerSerProAlaProGluPheLeuAlaSerSerProProApnIyS 878
 Db 2935 CTGACAGCTCTCGAGAGAGCGCTGACTATTCGCTGAGACAGGAAACAGACAG 2994
 Qy 878 rProApnGluAlaThrAlaAlaAp-----GlnIuserGluApnApnSerAl 895
 Db 2995 CTTGAACAGACAGTGAACGCGCGCGCTCCGCTGACTACGACGACGAAAGTTGACGCT 3054
 Qy 895 aSerArgThrSerLeuGluArgGlnAlaProHlArgGlyAenThrMetValHlIyValCy 915
 Db 3055 GCAGCAACAGCAGCAAAACAGACGACGACATCCGAAACATACGCGCTGACGCTCG 3114
 Qy 915 sTyRPhlArgApnThrSerValSerMetValApnPheSerIleAlaValGluApnIle 935
 Db 3115 TTGGCATCTGTGCGCCGACCGTGAAGTCCGATGACCTGTGAATAGCCGAGACACAGT 3174
 Qy 935 uSerGluApnLeuArgIyPheIyApnSerAenGlyTrpGlnIyLeuTrpValIa 955
 Db 3175 GTCCGGTATCTGCTGCGAAAGTTCAAGATAGCTCCGCTGACAGAGCTCTGGGTGT 3234
 Qy 955 lPheThrApnPheCybLeuPhePheTyRlySerSerHlAlaApnApnHlProleuAla 975
 Db 3235 GTTCACGCTCTTGTCTGTACTTTTACAAAGACTACAGACAGAGTTCGACGTGCTAG 3294
 Qy 975 rLeuProLeuLeuGlyTyRSerLeuThrlProSerGluSerGluApnIleGlnIyS 995
 Db 3295 TCTGCCGCTATTGGTTTACAGACGCTTCTCTGCTCCGCTGCTCAACGAAGCCGTTGAGAAGA 3354
 Qy 995 pTyRValPheIyLeuHlPheIySerHlIyValTyRTrpPheArgAlaGluSerGluTy 1015
 Db 3355 GTTCGCTTCAGACTTCTTTAAGAACACGCTACTTCTTCGCGGAGAAAGTCCGA 3414
 Qy 1015 rThrPheGluArgTrpMetGluValIleArgSerAlaThrSerSerAlaSerArgProHl 1035
 Db 3415 CACCTAACAAAGATGCTGAGAGTACTGCGACAGCAACCAACCCAGAGACTTCAAGAA 3474
 Qy 1035 sValLeuSerHlS 1039
 Db 3475 CGTACACAGCCAT 3487
 RESULT 15
 BC027077 2431 bp mRNA linear ROD 21-OCT-2003
 LOCUS Mus musculus FERM, RHOGEF (Arhgef) and plectrokin domain protein 1
 DEFINITION (chondrocyte-derived), mRNA (cDNA clone IMAGE:5010682), partial
 cds.
 ACCESSION BC027077
 VERSION BC027077.1 GI:20071584
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2431)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stetson M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedl T.B., Toohyuk S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McKean P.J., McKernan K.D., Malek U.A., Gunaratne P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fehey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Gitwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smalins D.E., Scherach A., Schein J.E., Jones S.J. and Marra M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2431)

Strausberg R.

Direct Submission

Submitted (04-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

Contact: MGC help desk

Email: gcgaps-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanaavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: IRAP Plate: 45 Row: C Column: 12

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Source

Location/Qualifiers

1..2431

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N-3"

/db_xref="taxon:10090"

/clone="IMAGE:5010682"

/cbase_type="mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."

/clone_id="NCI CGAP_Mam2"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6"

<1..2431

/gene="Farp1"

/note="Synonym: Cdep"

/db_xref="GeneID:223254"

/db_xref="MGI:2446173"

<1..1065

/gene="Farp1"

/codon_start=1

/product="Farp1 protein"

/protein_id="AAH27077.1"

/db_xref="GI:20071585"

/db_xref="GeneID:223254"

CDS

/db_xref="MGI:2446173"

/translation="YKVVLRRLCKHPNPNADPRDCAALAEITMVAOLGTMKME
NPKQLHLKQDLGIDNLTTPGRRFRLGSLKSGGLQQRNPLFNDLTLSRL
TASNQFVHQLDLYGTTTIESEBEWVPECLTLRGROSIIVAASRSMEKMEI
QMAIDLAENSGPTPELNASPPDNKSPDADADSEEDLSASRSLEKQPHRN
TQNHDAENRSTSVSMVDFSLAVENQSGNLRFRKSNMGOKLVFTNCLPEYKSH
QDHPPLASLPLIGSLTIPSESENIHVDYFKLHFKSHVYFPAESYTERMWEVIR
SARSSPASRIHLHSHKSHLY"

250..501

/gene="Farp1"

/note="PH: Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, sytrrophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations around the predicted binding site for phosphatidylinositol lipide"

/db_xref="CD:smart00233"

730..1020

/gene="Farp1"

/note="PH: Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, sytrrophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations around the predicted binding site for phosphatidylinositol lipide"

/db_xref="CD:smart00233"

misc_feature

misc_feature

ORIGIN

Alignment Scores:

Pred. No.: 1,54e-136 Length: 2431

Score: 1784.00 Matches: 338

Percent Similarity: 97.7% Conservative: 8

Best Local Similarity: 95.5% Mismatch: 8

Query Match: 32.7% Indels: 0

DB: Gaps: 0

US-09-555-342B-2 (1-1045) x BC027077 (1-2431)

692 TTTTGGTAAAGGATGCTGCTGAGAGCTGTCGACACACCAACCAACGCGCTTACAG

1 TATTAAGCATGCTCTGAGAGCTGTCGACACCAACCAACGCGCTTACAG 60

712 ATGCGAAGTGAAGAACTTCGAAAGCTGCAAGCTCAAGAAAGTCTGATCGGATGAC

61 GACTGCAAGAGCTGCTGCGAGATCACAGATGCTGACAGCTGATGATGATGATG 120

732 TTTTGGTAAAGGATGCTGCTGAGAGCTGTCGACACCAACCAACGCGCTTACAG

121 ATCAAGTGAAGAACTTCGAAAGCTGCAAGCTCAAGAAAGTCTGATCGGATGAC 180

752 ATGCGAAGTGAAGAACTTCGAAAGCTGCAAGCTCAAGAAAGTCTGATCGGATGAC

181 AATCTTGTGACCCAGAGAGGATTCATCGCTGCGGACCTCAGCAAGCTTCGCGG 240

772 TTTTGGTAAAGGATGCTGCTGAGAGCTGTCGACACCAACCAACGCGCTTACAG

241 AAGGAGCTTACAGAGCTGCTTCTTCTTCAAGATGCTTCTGCTGATACAGCGG 300

792 GATGCGAAGTGAAGAACTTCGAAAGCTGCAAGCTCAAGAAAGTCTGATCGGATGAC

301 GAACTGCAAGAGCTGCTGCGAGATCACAGATGCTGACAGCTGATGATGATGATGAC 360

```
QY 812 ILeGIuIuSerGIuApeGIuTrpGIuValProHieCySeruThrLeuArgGIuIuArg 831
DB 361 ATCGAGAGAGGTAGAGAGGTGGGTGTGCCCCACTGCTGACCTTGCGGGGCCAGCG 420
QY 832 GlnSerIleIleValAlaIaSerSerArgSerGIuMetGIuIuArgTrpValGIuApeIle 851
DB 421 CAATCTATCATCTGGGTGCTGAGTCCCGGTCAAGATGGAATAATGATGAGAGACATC 480
QY 852 GlnMetAlaIleApeLeuAlaGIuIuSerSerSerProAlaProGIuPheLeuAlaSer 871
DB 481 CAGATGCCCCATTGACCTGCGAGAGAAAGCAATGAGCCACCCCGAGCTCTTGCGAGC 540
QY 872 SerProProApeApeIuSerSerProApeGIuAlaThrAlaAlaApeGIuIuSerGIuApe 891
DB 541 AGCCCCCTTACACAAAGATCCCGGATGAGACCAAGCGGTGACCAAGAGTCAAGAGAT 600
QY 892 ApeIuSerAlaSerArgTrpSerLeuGIuArgGIuAlaProHieArgGIuApeThrMet 911
DB 601 GACCTGAGTGCCTCTCGGACATCGCTGAGCGCCAGGCCCTCACCGTGGCAACAGATG 660
QY 912 ValHieValCysTrpHisArgAenThrSerValSerMetValApePheSerIleAlaVal 931
DB 661 GTGCACGTGTGCTGGCATCGACACACCATGTCTCCATGTGTGACTTCAGCATTTGCTGTG 720
QY 932 GIuApeGIuIuSerGIuApeLeuLeuArgGIuApePheIuApeSerApeArgIuTrpGIuIuArg 951
DB 721 GAGAACCACTCTCTGGGAACTTGCTGAGAAAGTTCAAGAACAGCAATGCTGGCAGAG 780
QY 952 LeuTrpValValPheThrApePheCysLeuPhePheTyrIuSerHisGlnApeApeHis 971
DB 781 CTGTGGGTGTATTCACCACTTCTGCTGTCTTCTTCAAAATCACACGAGACAGTCAAC 840
QY 972 ProLeuAlaSerLeuProLeuLeuGIuTyrSerLeuThrIleProSerGIuSerGIuApe 991
DB 841 CCCCTTGCCAGCCTGCTGCTGGGCTACTCCCTCACCATCCCTCTGAGTGTGAGAAC 900
QY 992 ILeGIuIuApeTrpValPheIuSerHisPheIuSerHisValTyrTyrPheArgAla 1011
DB 901 ATCCACAGAGACTATGTGTTCAGAGCTGCAATTCAGTCCACGCTGACTTCAAGGCT 960
QY 1012 GIuSerGIuTyrThrPheGIuArgTrpMetGIuValIleArgSerAlaThrSerSerAla 1031
DB 961 GAAAGCGAATACACATTGAAAGTGGATGGAAGTATCCGAAGTGCACACAGCTCAGCC 1020
QY 1032 SerArgProHisValLeuSerHisIuSerGIuSerLeuValTyr 1045
DB 1021 TCTGGGCCCACTTGTAGTCAACAAATCTCACCTGTAC 1062
```

Search completed: May 20, 2006, 05:44:36
Job time : 14331 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 20, 2006, 01:25:21, Search time 1473 Seconds
(without alignments)
7419.552 Million cell updates/sec

Title: US-09-555-342B-2
Perfect score: 5463
Sequence: 1 MGIEBRPPTGSRGAPENS.....SATSSARPHVLSHKESLVY 1045

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5244920 seqs, 3486124231 residues
Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+.p2n.model -DEV=xlh
-Qc/abses/ABSSWB.spool/US09555342/runcat.19052006.121452.13173/app_query.fasta.1
-DB=N.GeneSeq -QMT=fastcap -SUPFIX=p2n.rng -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs05h
-USER=US09555342_QCEN_1.1.761@runcat.19052006.121452.13173 -NCPU=6 -ICPU=3
-NO_MMP -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N_GeneSeq.8:*

1:	geneeqn1980a:*
2:	geneeqn1990a:*
3:	geneeqn2000a:*
4:	geneeqn2001a:*
5:	geneeqn2001b:*
6:	geneeqn2002a:*
7:	geneeqn2002b:*
8:	geneeqn2003a:*
9:	geneeqn2003b:*
10:	geneeqn2003c:*
11:	geneeqn2003d:*
12:	geneeqn2004a:*
13:	geneeqn2004b:*
14:	geneeqn2005a:*
15:	geneeqn2006a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5463	100.0	3442	2	AAx79183
2	5463	100.0	3442	15	AEEO3370
3	5463	100.0	5028	14	AEb22714

4	5459	99.9	4687	3	AAA08582	aa08582 Human cyt
5	5431	99.4	3187	5	AA64760	aa64760 DNA encod
6	2941.5	53.8	3997	12	ADQ84807	adq84807 Human tun
7	2941.5	53.8	3997	13	ADR25675	adr25675 Breast ca
8	2941.5	53.8	3997	13	AD644452	ad644452 Human KIA
9	2498	45.7	3094	3	AAc98992	aac98992 Human pan
10	1150.5	21.1	3200	12	ADOC6863	adoc6863 Novel hum
11	1006.5	18.4	1718	4	AA527016	aa527016 cDNA enco
12	1006.5	18.4	1718	10	ADB93194	adb93194 Human cDN
13	964	17.6	582	10	ACD92278	acd92278 Human col
14	964	17.6	582	10	ACD94832	acd94832 Human col
15	913.5	16.7	1747	4	ABL03867	abl03867 Drosophi1
16	905.5	16.6	1220	4	ABL02335	abl02335 Drosophi1
17	891.5	16.3	4150	4	ABL03866	abl03866 Drosophi1
18	754.5	13.8	2277	10	ADCT1321	adct1321 Human col
19	754.5	13.8	3275	13	ADRO7347	adro7347 Full leng
20	748.5	13.7	3309	13	ADRI4498	adri4498 Human NF-
21	748.5	13.7	3309	14	ADY16637	ady16637 DNA encod
22	748.5	13.7	3309	14	ADY15739	ady15739 DNA encod
23	748.5	13.7	3309	14	AD222659	ad222659 Erythrocy
24	744	13.6	4543	10	ABT42346	abt42346 Toxicity
25	743.5	13.6	3387	13	ADRI4396	adri4396 Human NF-
26	739.5	13.5	3770	13	ADRO8260	adro8260 Full leng
27	735.5	13.5	14917	6	AA032338	aa032338 Human lun
28	727	13.3	6305	10	ADC31979	adc31979 Human nov
29	727	13.3	6305	10	ADC31980	adc31980 Human nov
30	724	13.3	3620	3	AA08581	aa08581 Human cye
31	724	13.3	3620	10	ADJ56474	adj56474 Murine CD
32	724	13.3	4336	10	ADH28737	adh28737 Human chr
33	724	13.3	4336	14	ADY14625	ady14625 DNA encod
34	724	13.3	4336	14	ADY14627	ady14627 DNA encod
35	724	13.3	4336	14	ABG82823	abg82823 Breast ca
36	717.5	13.1	6274	10	ADC30054	adc30054 Human nov
37	717	13.1	637	14	ACLS5653	ac15653 Human col
38	716.5	13.1	2410	8	ACC44336	acc44336 Gene enco
39	716.5	13.1	4338	14	AEA19657	aea19657 Novel hum
40	715	13.1	671	13	ADQ52709	adq52709 Novel can
41	714	13.1	435	14	ACI61462	aci61462 Human col
42	713	13.1	2656	14	AEA19659	aea19659 Novel hum
43	713	13.1	2867	12	ADP10335	adp10335 Reference
44	713	13.1	2867	13	ADR90365	adr90365 Human ful
45	713	13.1	2867	13	ADP54688	adp54688 Human PRO

ALIGNMENTS

RESULT 1	AAx79183	standard; DNA; 3442 BP.
ID	AAx79183	
XX	AAx79183;	
AC	17-AUG-1999	(first entry)
DT	Human chondrocyte-derived gene CDEP.	
XX		
DE		
XX		
KW	Differentiation; human; foetal; chondrocyte; ezrin-like domain; cancer;	
KW	DBI homology domain; plectstrin homology domain; rheumatoid 'arthritis;	
KW	drug; ss.	
XX		
OS	Homo sapiens.	
XX		
XX	WO9928458-A1.	
XX	10-JUN-1999.	
PD		
XX		
PF	27-NOV-1998;	98MO-JP005348.
XX		
PR	27-NOV-1997;	97DP-00342060.
XX		
PA	(CHUS) CHUGAI SEIYAKU KK.	
XX		
PI	Kato Y, Kawamoto T, Koyano Y;	

XX WPI: 1999-371117/31.
DR P-PSDB; AA07482.
XX Protein CDEP expressed in differentiated chondrocytes, and gene encoding
PT it.
XX
XX Claim 5; Fig 1; 59pp; Japanese.

CC This sequence represents the coding region for a protein (CDEP) expressed
CC in differentiated human foetal chondrocytes, which contains an ezrin-like
CC domain, a Dbl homology (DH) domain and a pleckstrin homology (PH) domain.
CC The nucleic acid or protein can be used in the investigation and
CC treatment of cancers and arthritic diseases (including chronic rheumatoid
CC arthritis), or for screening of candidate anticancer drugs

XX Sequence 3442 BP; 864 A; 952 C; 927 G; 699 T; 0 U; 0 Other;

Alignment Scores:

	0	Length:	3442
Pred. No.:	5463.00	Matches:	1045
Score:	100.0%	Conservative:	0
Percent Similarity:	100.0%	Mismatches:	0
Best Local Similarity:	100.0%	Indels:	0
Query Match:	2	Gaps:	0

US-09-555-342B-2 (1-1045) x AAX79183 (1-3442)

QY 1 MercGlyGluIleGluGlnArgProThrProGlySerArgLeuGluValAProGluAAsnSer 20
DB 49 ATGGGAAATATAGACAGAGCCGACCCCGATCAAGACTGGGGGCCCGGAAATTCG 108
QY 21 GLLIleSerThrLeuGluArgGlyGlnIleValProProThrProSerGlyLysLeuVal 40
DB 109 GGGATGATGACCTTGGACGTGGACAGACCCGCCCAACCTTCAGAAATCTGTG 168
QY 41 SerIleLysIleGlnMetLeuAspThrGlnGluValPheGluValProGlnArgAla 60
DB 169 TCATCAAAATCCAGATGCTGGATGACACCCAGAGGCATTGAAAGTTCCAAAGAGCT 228
QY 61 ProGlyLysValIleuLeuAspAlaValCysAsnHisLeuAsnIleuValGluGlyAspTyr 80
DB 229 CCGGGAGAGGTGCTGCGATGACATTTGGACACCACTCAACCTCGGGAGGTGACTAT 288
QY 81 PheGlyLeuGluIlePheProAspHisLysLysIleThrValTyrLeuAspIleuLeuLysPro 100
DB 289 TTGGGCTCGAGTTTCTCTGATCAAAAAGATCAAGGTGGCTGATCTCTAAAACCC 348
QY 101 IleValLysGlnIleArgArgProLysHisValValIleAspPheValValLysPhePhe 120
DB 349 ATGTGAAACAGATTAGAAAGCCAAAGCAGTTGTTGTAAAGTTTGGTGAATTCCTT 408
QY 121 ProProAspHisThrGlnLeuGlnGluIleuLeuThrArgTyrIleuPheAlaLeuGlnVal 140
DB 409 CCCCCTGACACACAACTCCAAAGAACTCAAGAGTACCTGTTCGGCTGCAGGTG 468
QY 141 LysGlyAspLeuIleGlnGlyArgLeuThrCysAsnAspThrSerAlaIleuLeuIle 160
DB 469 AAGCAGAGCTTGGCTCAAGGAGGTTGACGTGTAATGACACAGCGGACTCTCTTGATT 528
QY 161 SerHisIleValGlnSerGluIleGlyAspPheAspGluValIleuAspArgGluHisLeu 180
DB 529 TCACACTGTTGCAATCTGAGATTGGGGATTTTGATGATGACCTTGGACAGAGCACTTA 588
QY 181 AlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIleValGluPheHis 200
DB 589 GCAAAAATTAATACATCTCAGCAGACAGCACTAAGAGCAAAATCGTGAATTCAC 648
QY 201 HisAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnIleuLeuGluIleAlaArg 220
DB 649 CATTAACCACTTGGACAAACACAGCAGATCAGATTTCCAGCTCTTGAAGATTGGCCGT 708
QY 221 ArgLeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysIle 240

DB 709 CGGCTAGAGATGATATGAAATCCGGTTGCACCCGGCCAAAGACAGGAGCAAGCAAGATC 768
QY 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAla 260
DB 769 AATCTGGCCGTTGCCCAACCGGGAATCTAAGTGTTCAGGGTTTCATAGATCAATATGCC 828
QY 261 PheAsnTrpAlaLysValArgLysLeuSerPheLysArgLysPheLeuIleLysLeu 280
DB 829 TTCACTGGGCCAAGGTGGGAAGCTGAGCTTCAAGAGAAAGGCTTCTTCATCAAGCTC 888
QY 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetLaseArgAsp 300
DB 889 CGGCCAGATGCCAATATGTGCTGATCAGATACCTTGGAATTCCTGATGGCCAGTCGGAT 948
QY 301 PheCysLysSerPheThrLysIleCysValGluHisAlaPhePheArgLeuPheGlu 320
DB 949 TTCTGCAATCTCTTCGGAATAATCTGTGTAACATCATAGCTTCTTGAATTTTGA 1008
QY 321 GluProLysProLysProLysProValLeuPheSerArgGlySerSerPheArgPheSer 340
DB 1009 GAGCCCAACCAAGCCCAAGCCCGTCTCTTACCGGGGGTCATCATTTCCGTTCAAT 1068
QY 341 GlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyGlyHisLysLysValGln 360
DB 1069 GGTCCGACTCAGAGACAGGTTCTCGACTATGTTAAAGAGAGACATAAAGAGTGCAG 1128
QY 361 PheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
DB 1129 TTGAAAGAGACACGAAAGATCTATCTACCGAGCTTGCTTCAAGCTTACAGAA 1188
QY 381 LeuAsnSerGluValLeuGlnIleSerGlnIleSerThrSerLeuThrPheGlyGluGly 400
DB 1189 CTGAATTCGGAAGTGTGAGAGAGTCTCAGCAGAGACCAAGCTTCAATTTGGAGAAAGT 1248
QY 401 AlaGluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGly 420
DB 1249 GCCGAATCTCAGAGGGGCGAGAGCTGCCGCGAGAAAGAACCGAAGGTTTCCGGCG 1308
QY 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAla 440
DB 1309 GAGCCGGGTCCGACCCGAGCCCTGCGCAGAGAGAAAGCCCGGGGTAAACAAGCAGCG 1368
QY 441 AspGlyAlaIleSerIleProThrGlnGluGlnGluGluValValIleAspArgThrGln 460
DB 1369 GACGAGCGCGCTCGCGCCACGAGAGAAAGAGAGAGTGTAAAGATAGAACCCAG 1428
QY 461 GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeu 480
DB 1429 CAGAGTAAACCTCAGCCCGCAGCCMACACAGAGCTCCCTGACTGCGAGTCTCACTT 1488
QY 481 SerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro 500
DB 1489 TCCAGACTGTCTGTGAATCTGCAAGGGGAGATGGCCCTTGCAACCTGATCTTCCC 1548
QY 501 AsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuLeuAsnAspGln 520
DB 1549 AACCTGAGCCCGACACCAAGAGAGGCTCTCCCTTATCAGCCCGGTGATGAGCAAG 1608
QY 521 AlaCysProArgThrAspAspGluAspGluGlyArgLysArgPheProThrAspLys 540
DB 1609 GCCTGCCCCCGACGACATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATTA 1668
QY 541 AlaTyrPheIleAlaLysGluValSerThrThrGlnLysThrTyrLeuLysAspLeuGlu 560
DB 1669 GCGTACTTAATAGCTAAGAAAGTGTACCAACCGCAACATATTTGAAGATCTCGAA 1728
QY 561 ValIleThrSerTyrPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeu 580
DB 1729 GTTATACATCTTGGTGTTCAGAGCAGTGAAGAAAGAGAGCGCATTCGCGAAGCAGCTG 1788
QY 581 LysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAsnPheLeuLys 600

Db 1789 AAAAGTCATATATCCGAAATTTTGAACCTTTGGACAAATTTACTAATTTTCTCAG 1848
 Qy 601 G|U|I|E|G|U|G|N|A|G|U|A|A|L|E|U|T|P|G|U|G|Y|A|T|S|E|R|A|N|A|A|G|I|N|I|E|A|T|A|S|P|T|Y|R 620
 Db 1849 GAAATTAGCAACACATTTGCTTGGAGAGCGCTCAATATGCTCAATATGAGATTAAC 1908
 Qy 621 G|A|T|A|T|G|I|E|G|Y|A|P|V|A|I|E|L|E|U|T|Y|E|A|N|I|E|G|I|N|G|Y|E|L|Y|S|E|R|S|E|R|A|A|I|A|H|S 640
 Db 1909 CAAAGATCCGGCGATGTCATCTCAAGAACATTTACAGGCGATGAAGACCTGGCCGCTCAC 1968
 Qy 641 L|E|U|T|P|Y|H|S|E|R|S|E|R|G|U|A|I|E|U|G|U|A|I|E|U|G|U|A|N|G|Y|I|E|Y|S|E|R|S|E|R|A|A|T|G 660
 Db 1969 CTGTGGAAGCACACAGAGCGCTTGGAGGCTTGGAGAAATGGAATCAAGAGCTTCCCGCGG 2028
 Qy 661 L|E|U|G|U|A|N|P|H|E|C|Y|A|T|A|S|P|H|E|G|U|E|U|G|I|N|Y|S|A|I|C|Y|S|T|Y|R|L|E|U|P|L|E|U|A|N|T|H|R 680
 Db 2029 CTGGAGAACTTCTCAGAGACTTGAAGCTGCAAGAGGTGTCTCAACCTCAACACC 2088
 Qy 681 P|H|E|U|E|U|A|R|P|R|O|L|E|U|H|S|A|R|G|L|E|U|E|H|S|T|Y|R|Y|S|G|I|N|V|A|I|L|E|U|G|U|A|R|G|L|E|U|C|Y|S 700
 Db 2089 TTCTCTCTGGGGCCACCTGACACCGGCTCATGCACTACAGAGGTCCTGGAGCGGCTGTGC 2148
 Qy 701 L|Y|E|H|S|H|S|P|R|O|S|E|R|H|S|A|I|A|A|P|H|E|A|R|G|A|P|C|Y|A|T|A|A|I|A|L|E|U|A|I|G|U|I|E 720
 Db 2149 AAACACCAACCGCGGAGCCGACGCTTCAAGGAGCTGCGAGCGCTTGGCAGAGATC 2208
 Qy 721 T|H|R|G|U|E|T|V|A|I|A|G|I|N|E|U|H|S|G|I|Y|T|H|E|T|I|E|Y|E|T|G|U|A|N|P|H|E|G|I|N|Y|L|E|U 740
 Db 2209 ACCGAGATGTGTGACAGCTCCACGCTACGATGATCAAGATGAGAAATTTCTCAAGACGTG 2268
 Qy 741 H|I|S|G|U|E|U|Y|S|Y|A|P|L|E|U|I|E|G|Y|I|L|E|A|P|H|E|U|V|A|I|P|R|O|G|I|A|R|G|U|P|H|E 760
 Db 2269 CACGAATCTCAGAAAGATTTTATTTGCACTTATGACATCTTGTGCTCCGGAAAGGAGATTC 2328
 Qy 761 I|L|E|A|R|G|E|U|G|Y|S|E|R|S|E|R|Y|E|U|S|E|R|G|Y|L|Y|S|G|I|N|G|I|N|A|R|G|E|U|P|H|E 780
 Db 2329 ATCCGTCTGGGGCAGCTCAGCAAGCTCTCGGGAAAGGGGCTCCAGCAGCGCATGTTCTTC 2388
 Qy 781 L|E|U|P|H|E|A|N|P|V|A|I|L|E|U|E|U|Y|R|T|H|S|E|R|A|R|G|I|Y|L|E|U|T|H|R|A|S|E|R|A|N|G|I|N|P|H|E|Y|S 800
 Db 2389 CTGTTCACACGACGCTCGCTATACACAGCGCGGGGCTGAGAGGCTCCAACTCACTTTAAA 2448
 Qy 801 V|A|I|H|S|G|I|N|E|U|P|R|O|L|E|U|Y|R|G|I|Y|E|T|H|R|I|E|G|U|G|U|S|E|R|G|U|A|P|R|G|I|T|T|R|G|I|Y 820
 Db 2449 GTCCACGGGACGCTCCGCTTATGCACTGACATTTAGAGAGACGAAAGACGATGGGG 2508
 Qy 821 V|A|P|R|O|H|S|C|Y|L|E|U|T|H|R|L|E|U|A|R|G|I|Y|G|I|N|A|R|G|I|N|S|E|R|I|E|I|V|A|I|A|I|S|E|R|S|E|R 840
 Db 2509 GTGCCCACTGCTGACCTCCGCGGGCCAGCGGCAGTCCATCATCTGTGGCCGACATTC 2568
 Qy 841 A|R|G|S|E|R|G|U|E|C|G|U|Y|T|T|R|V|A|I|G|U|A|P|R|I|E|G|I|N|E|C|A|I|I|E|A|P|L|E|U|A|I|G|U|Y|S 860
 Db 2569 CGGTCCAGATGAGAGATGGGTGAGACATCCAGATGGCCATTTGACCTGGCCGAGAGAG 2628
 Qy 861 S|E|R|S|E|R|P|R|O|A|I|P|R|O|G|I|N|P|H|E|U|A|I|S|E|R|S|E|R|P|R|O|A|P|A|N|Y|S|E|R|P|R|O|A|S|P 880
 Db 2629 AGCGACGAGCCCGCTGATCTGCTGCGCAGAGCCCGCTGCAACAAATCTCCCTGAT 2688
 Qy 881 G|U|A|I|A|T|H|R|A|A|A|A|P|R|G|I|N|S|E|R|G|U|A|P|A|P|L|E|U|S|E|R|A|I|S|E|R|A|R|G|I|T|H|S|E|R|U 900
 Db 2689 GAAGCCACCGCGGCTGACAGAGATCAAGAGATGACCGGCTGCGGACATCCGCTG 2748
 Qy 901 G|U|A|R|G|I|N|A|P|R|O|H|S|A|R|G|I|Y|A|N|T|H|E|T|V|A|I|H|S|V|A|I|C|Y|S|T|T|R|P|H|S|A|R|A|N|T|H|R 920
 Db 2749 GAGGCGAGGCGCGGACCGCGGCAACAAATGTGTGACGCTGTCTGCGACCGGAAACACC 2808
 Qy 921 S|E|R|V|A|I|S|E|R|E|T|V|A|I|A|P|H|S|E|R|I|E|I|A|I|G|U|A|N|G|I|N|S|E|R|G|I|Y|A|N|L|E|U|U 940
 Db 2809 AGGCTCTCCATGTGTGACTTACAGCATCGCATGAGATCGATGCTGTCTGAAACCTGCTG 2868
 Qy 941 A|R|G|Y|A|P|H|E|Y|A|A|N|S|E|R|A|N|G|I|Y|T|R|P|G|I|N|Y|S|E|R|U|P|V|A|I|P|H|E|R|A|N|P|H|E|C|Y|S 960
 Db 2869 AGGAAATTCAAAACACAGACGCGGTGGCAGAAAGCTGTGGGTGTTCACAAATCTTCTGC 2928

Qy 961 L|E|U|P|H|E|H|E|R|Y|R|S|E|R|H|S|G|I|N|A|P|A|N|H|S|P|R|O|L|E|U|A|I|S|E|R|L|E|U|P|R|O|L|E|U|G|Y 980
 Db 2929 CTGTCTTCTTCAAAATACACAGAGACAAATATATCCCTTGGCAACCTGCTTGTCTGGC 2988
 Qy 981 T|Y|S|E|R|L|E|U|T|H|R|I|E|P|R|O|S|E|R|G|U|S|E|R|G|U|A|N|I|E|G|I|N|Y|S|A|P|T|Y|R|V|A|I|P|H|E|Y|S|E|R 1000
 Db 2989 TACTCGTCAACATCCCTCTGAGTCCGAGAACATCCAGAAAGACTACGTTCTCAACTG 3048
 Qy 1001 H|I|S|P|H|E|Y|S|E|R|H|S|V|A|I|Y|R|Y|R|P|H|E|A|R|G|I|A|G|U|S|E|R|G|U|Y|R|T|H|R|P|H|E|G|U|A|R|T|P 1020
 Db 3049 CACTTCAAGTCCCACTGCTACTACTTCAAGCGCGGAAAGCCAGATACAGTTCGAAAGGTGG 3108
 Qy 1021 M|E|R|G|U|V|A|I|I|E|A|R|S|E|R|A|I|A|T|H|S|E|R|S|E|R|A|S|E|R|A|R|P|R|O|H|S|V|A|I|L|E|U|S|E|R|H|S|Y|S 1040
 Db 3109 ATGGAAGTATCCGCAAGTGCACACGCTGTGCTTGGCAGCCCGAGTGTGAGCCACAAA 3168
 Qy 1041 G|U|S|E|R|L|E|U|V|A|I|Y|R 1045
 Db 3169 GAGTCTTGTGTAT 3183
 RESULT 2
 ABE03370 ID ABE03370 standard; DNA; 3442 BP.
 AC ABE03370;
 DT 09-FEB-2006 (first entry)
 DE XX
 KW Acute myelogenous leukemia prognosis related DNA sequence SEQ ID NO: 31.
 KW dg; gene; acute myelogenous leukemia; prognosis; gene expression;
 KW biochip.
 XX Homo sapiens.
 OS JP200533987-A.
 PN 08-DEC-2005.
 PD 06-MAY-2005; 2005JP-00135284.
 PF 06-MAY-2004; 2004US-0568635P.
 PR (VERT-) VBRIDEX LLC.
 PA Lapont M;
 PI WPI; 2006-004067/01.
 DR
 XX
 PT Evaluating an acute myelogenous leukemia patient's prognosis, comprising
 PT detecting lower or higher expression level of gene recognized by probe
 PT set e.g. 202820-at and 206148-at, with respect to predetermined cut-off
 PT level.
 PS
 XX Example 5; SEQ ID NO 31; 60bp; Japanese.
 XX
 CC The present sequence is that of a human gene sequence which is claimed
 CC for use in evaluating the prognosis of patients suffering from acute
 CC myelogenous leukemia (AML) by analysis of the patients gene expression
 CC profile. The invention relates to a novel method for evaluating the
 CC prognosis of a patient with acute myelogenous leukemia by detecting a
 CC high/low expression level of genes encoding mRNA recognized by a
 CC probe set chosen from 19 nucleotide sequences (ABE03340 or ABE03344-
 CC ABE03361). Also, gene expression profiling of AML patients using a set of
 CC probes based on 167 sequences (ABE03362-ABE03527) using a biochip was
 CC performed to determine prognosis. Methods are also included for
 CC determining an AML patient's state, determining AML patient's treatment
 CC protocol, determining whether the patient will respond to the treatment,
 CC and producing an AML patient's prognosis report by analysis of the
 CC patients gene expression profile. The methods are useful in evaluating an
 CC AML patient's prognosis and for treating AML patient.

Sequence 3442 BP; 864 A; 952 C; 927 G; 699 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3442
Score:	5463.00	Matches:	1045
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	15	Indels:	0
		Gaps:	0

US-09-555-342B-2 (1-1045) x AEE03370 (1-3442)

QY 1 MetGlyGluIleGluGlnArgProThrProGlySerArgLeuGlyAlaProGluAsnSer 20
DB 49 ATGGAGAAATTAAGACAGAGGCGGAGCCCGAGATCAGACTGGGGGCGCCGGAAAAATTCG 108
QY 21 GlyTLeSerThrLeuGluArgGlyGlnLysProProProThrProSerGlyLysLeuVal 40
DB 109 GGGATCAGTACTTGGAAACGTGGACAGAACCGCCCCAACACCTTCAGGAAAACTCGTG 168
QY 41 SerTLeuValLeuMetLeuAspThrGlnGluAlaPheGluValProGlnArgAla 60
DB 169 TCATCAAAATCCAGATGCTGAGATGACACCCAGAGGCAATTGAAGTTCCAAAGAGCT 228
QY 61 ProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGluGlyAspTyr 80
DB 229 CTGGGAGAGGTGCTGCTGAGATGCAAGTTTGCACACCTCAACCTCGTGGAGAGGACTAT 288
QY 81 PheGlyLeuGluPheProAspHisLysLysValLeuValTyrLeuAspLeuLeuLysPro 100
DB 289 TTTGGCTCAGATTCTCGATTCACAAAAAGATCAGGTGTGCTGAGATCTCTAAACC 348
QY 101 TLeValLysGlnTLeaArgProLysHisValValLysPheValLysPhePhe 120
DB 349 ATTTGAAACAGATTAGAGGCCAAGACACGTTGTATTAAAGTTTGTGTGAATTTCTT 408
QY 121 ProProAspHisThrGlnLeuGlnGluGlnLysThrArgTyrLeuPheAlaLeuGlnVal 140
DB 409 CCGGCTACCAACACACAACTCCAAGAAAGTCAACAAGTACCTGTGCGGTGCAAGGTG 468
QY 141 LysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAlaAlaLeuLeuTLe 160
DB 469 AAGGAGACATTGGCTCAAGCAGAGGTGACGTATATACACCAAGCGACGCTCTTGATT 528
QY 161 SerHisTLeValGlnSerGluTLeGlyAspPheAspGluAlaLeuAspAspGlnHisLeu 180
DB 529 TCACACATTGTGCAATGTGAATTTGGGAAATTTGATGAAGCTTGGACAGAGACACTTA 588
QY 181 AlaLysAsnLysTyrTLeProGlnGlnAspAlaLeuGluAspLysTLeValGluPheHis 200
DB 589 GCAAAAAATTAATATCATCACTCAGCAAGACGCACTAAGGCAAAATCGTGAATTTTCA 648
QY 201 HisAsnHisTLeGlyGlnThrProAlaGluSerAspPheGlnLeuLeuGlnTLeAlaArg 220
DB 649 CATAACCACTTGGACAAACACCAAGCAAGATTCAGATTTCCAGCTCTCAAGATTTGCCGT 708
QY 221 ArgLeuGlnLeuSerTyrGlyTLeArgLeuHisProAlaLysAspArgGluTyrLysTLe 240
DB 709 CGGCTACAGATGTAATGGAATCCGGTTGCACCCGCGCAAGACAGGAGGACAGAGATC 768
QY 241 AsnLeuAlaValAlaAsnThrGlyTLeLeuValPheGlnGlyPheThrLysTLeAsnAla 260
DB 769 AATCTGGCGTTGGCAACAGGGAAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 828
QY 261 PheAsnThrAlaLysValAlaGlyLysLeuSerPheLysArgLysArgPheLeuTLeLysLeu 280
DB 829 TTCAACTGGGCGCAAGGTGGAGAGCTGAGCTTCAAGAGGAAGCGCTTTTCATCAACAGCTC 888
QY 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAlaSerArgAsp 300
DB 889 GGGCGAATGCCAATAGTGGGTACCAAGATACCTTGAATTTCTGATGGCCAGTCCGGGAT 948
QY 301 PheCysLysSerPheThrLysTLeCysValGluHisAlaPhePheArgLeuPheGlu 320

DB 949 TTCTGCAAGTCCTTGGAAAAATCTGTGTTGAACATCAGCTCTTTAGACTTTTGAA 1008
QY 321 GluProLysProLysProLysProValLeuPheSerArgGlySerSerPheArgPheSer 340
DB 1009 GAAGCCCAACCAAAAGGCCAAGCCGTCCTTTTAAAGCGGGGGTCAATTCATTTGGTTCAGT 1068
QY 341 GlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyLysLysValGln 360
DB 1069 GGTGGAATCAGAGAGAGGTTCTCGACTATGTTTAAAGAAAGAGACATAAGAGCTGCAG 1128
QY 361 PheGluArgLysHisSerLysTLeHisSerTLeArgSerLeuAlaSerGlnProThrGlu 380
DB 1129 TTTGAAGAAGACACAGCAAGATTCATCTATCCGAGCCTTGCTTACAGCTTACAGAA 1188
QY 381 LeuAsnSerGluValLeuGlnGlnSerGlnGlnSerThrSerLeuThrPheGlyGly 400
DB 1189 CTGAATTTGGAAATGCTGGAGAGCTCTCAGACAGACACAGCTTACATTTGAGAAAGT 1248
QY 401 AlaGluSerProGlyLysGlnSerCysArgArgGlyLysGluProLysValSerAlaGly 420
DB 1249 GCCGAATCTCCAGAGGGGCGAGAGCTCCGGCGAGAAAGAACCGAAGTTTCCGCGGG 1308
QY 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAla 440
DB 1309 GAAGCGGGGTGCAACCGAGCTCGCGCGAGAGAGAAAGAGAGTCTTAAAGATAGACCCAG 1428
QY 441 AspGlyAlaAlaSerAlaProThrGlnGlnGlnGlnGlnValValLysAspArgThrGln 460
DB 1369 GAAGAGAGCGGCTCGGCGCCACGAGAGAAAGAGAGTCTTAAAGATAGACCCAG 1428
QY 461 GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeu 480
DB 1429 CAGAGTAACTCAGGCCCGCCGACGCCACAGACAGCTCTCTGACTGCACTTCACTT 1488
QY 481 SerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro 500
DB 1489 TTCGAGCTGTGTGAATCTGCAGAGGGGAGTGGCCCTGCGCAACCTGTGCTGCC 1548
QY 501 AsnLeuSerProAspThrLysGlnAlaSerProLeuTLeSerProLeuAsnAspGln 520
DB 1549 AACCTAAGCCCCACACCAAGAGGCTCTCTCTTATCAAGCCGCTGCTGAATTCAG 1608
QY 521 AlaCysProArgThrAspAspGluAspGluGlyArgArgLysArgPheProThrAspLys 540
DB 1609 GCCTGCCCGGACGACGATGAGATGAGGCGCGGAGAAAGATTTCCCACTGATTA 1668
QY 541 AlaTyrPheTLeAlaLysGluValSerThrThrGlnArgThrTyrLeuLysAspLeuGlu 560
DB 1669 GCGTACTTCATAGCTAAGAGAGTGTACCAACGAGCGAAATATCTGAAGATTCGAA 1728
QY 561 ValTLeThrSerThrPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeu 580
DB 1729 GTTATCATCTCGTGTTCAGAGACAGATGAGAAAGAGACCCATGCGGAGGACTG 1788
QY 581 LysSerLeuTLePheProAspPheGluProLeuHisLysPheHisThrAsnPheLeuLys 600
DB 1789 AAAAGTCTAATNTCCCGAATTTTGAACCTTTGCAAAATTTCAATGATTTTTCGAA 1848
QY 601 GluTLeGluGlnArgLeuAlaLeuTyrGluGlyArgSerAsnAlaGlnTLeArgAspTyr 620
DB 1849 GAAATTGACCAACGACTTCCCTGTGGAGAGCGCTCAAAATCCCAAAATCAAGATTC 1908
QY 621 GlnArgTLeGlyAspValMetLeuLysAsnTLeGlnGlyMetLysHisLeuAlaHis 640
DB 1909 CAAAGATGCGCATATCATGTGAAGAACTTCAAGGACATGAAGACCTCGGGGCTCAC 1968
QY 641 LeuTyrLysHisSerGlnAlaLeuGlnAlaLeuGluAsnGlyTLeLysSerSerArgArg 660
DB 1969 CTGTGGAACACACAGGCTTTGAGAGCTCTGGAAGATCGAATCAAGAGCTCCGGCG 2028
QY 661 LeuGluAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuProLeuAsnThr 680


```

Db      2029 CTGGAGAACTTCTGACAGACTTTGAGCTGCGAAGAGTGTTGTTACTACCGCTCAACACC 2088
Qy      681 PheLeuLeuAaGrProLeuHiaArgLeuMetHsTyrTyrGlnValLeuGluArgLeuCys 700
Db      2089 TTCTCTCGCGGACCATGACACCGGCTCATGCATCAAGACAGGCTCTTGAGCGGCTGGC 2148
Qy      701 LysHsHsAPrProSerHsEaLAspPheArgAspCysAspAlaAlaLeuAlaGluIle 720
Db      2149 AAACACCAACCCCGCCGACGCGACCGCATTCAGGAGCTCCAGACCGCTCTTGGCAGAGATC 2208
Qy      721 ThrglMetValAlaGlnLeuHsEglYThrMetIleLysMetGluAsnPheGlnLysLeu 740
Db      2209 ACGGAGATGATGGCAGACTCCACGGTACGATGATCAAGATGAGAAATTCACAGAAAGCTG 2268
Qy      741 HsEglLeuLysLysLysPheLeuIleGlyLLeaPheLeuValValProGlyValArgGluPhe 760
Db      2269 CACCAAACTCAAGAAAGATTATTTGGCATTCATCTTGGTTCGGGAAAGGAGATTC 2328
Qy      761 HLeArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArgMetPhePhe 780
Db      2329 ATCCGCTCGGACAGCTTCAGCAAGCTCTCGGAGAAAGGAGCTCCAGACCGCATGTTCTTC 2388
Qy      781 LeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsnGlnPheLys 800
Db      2389 CTGTTCAACGACGCTCTGTCTATACACGAGCCGGGGGCTGACGGCTCCAAATCAAGTTTAA 2448
Qy      801 ValHsEglGlnLeuProLeuTyrGlyMetThrIleGluGluSerGluAspGluTyrGly 820
Db      2449 GTCCACGGGAGCTCCCGCTTAAGCAATGACATGAGAGACGAGAGACGAGTGGGGG 2508
Qy      821 ValProHsCysLysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaLAserSer 840
Db      2509 GTGCCCACTGCTGACCTCCGGGGCCAGCGGACATCATCTGTGGCCGCAATTTCT 2568
Qy      841 ArgSerGluMetGluLysTyrValGluAspIleGlnMetAlaIleAspLeuAlaGluLys 860
Db      2569 CGGTCGAGATGAGAAAGTGGTGAAGGACATCCAGATGGCCATTTGACCTGCGGAGAAAG 2628
Qy      861 SerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLysSerProAsp 880
Db      2629 AGCAGCAGCCCCCGCTGATCTCTGACGACGACCCCCCTGCAACAACTCCCTGAT 2688
Qy      881 GluAlaThrAlaAlaAspGlnGluSerGluAspAspLeuSerAlaSerArgThrSerLeu 900
Db      2689 GAAACCAACCGGCTGACACAGACTCAGAGATTAACCTGACGCTCGCGCAATCGCTG 2748
Qy      901 GluArgGlnAlaProHsArgGlyAsnThrMetValHsValCysTyrPHisArgAsnThr 920
Db      2749 GAGGCGCAGGCCCCGACACCGGCAACAAATGGTGGACGTTGCTGGCACCGCAACACC 2808
Qy      921 SerValSerMetValaAspPheSerIleAlaValGluAsnGlnLeuSerGlyAsnLeuLeu 940
Db      2809 AGCCTCTCCAGTGGGACTTCAGCATCGCAGTGAAGATCAAGTGTCTGGAAACCTGTG 2868
Qy      941 ArgLysPheLysAsnSerAsnGlyTyrGlnLysLeuTyrPValIlePheHisAsnPheCys 960
Db      2869 AGGAAATTCAAAACACGACGAGCGGTGGCAAAAGCTGTGGGTTCACAAACTTCTGC 2928
Qy      961 LeuPhePheTyrLysSerHsEglAspAsnHsIaProLeuAlaSerLeuProLeuGly 980
Db      2929 CTGTTCTTTCACAAATCACACAGCAATATCATCCCTTGGCAGCTCGCTCTGCTGGC 2988
Qy      981 TyrSerLeuThrIleProSerGlnSerGluAsnIleGlnLysAspTyrValPheLysLeu 1000
Db      2989 TACTCGCTCAACATCCCTCGAGTTCGAGAAACATCCAGAAAGCTACGTTTCAAGCTG 3048
Qy      1001 HsPheLysSerHsValIlyTyrPheArgAlaGluSerGluTyrThrPheGluArgTyr 1020
Db      3049 CACTTCAAGTCCACGTTTACTTACTTACGGCGGAAAGCAGATCAAGCTTGAAGGTGG 3108
Qy      1021 MetGluValIleArgSerAlaThrSerSerAlaSerArgProHsValIleLeuSerHsLys 1040
Db      3109 ATGGAAAGTGAATCCGACAGTGCACACAGCTCTGCTCGGCAACCCACGTTTGAAGCCACAAA 3168

```

```

Qy      1041 GluSerLeuValTyr 1045
Db      3169 GAGTCTCTGTGTAT 3183

RESULT 3
AEB22714
ID      AEB22714 standard; DNA; 5028 BP.
XX
AC      AEB22714;
XX
DT      22-SEP-2005 (first entry)
XX
DE      Human colon cancer-associated gene SEQ ID NO:9.
XX
KW      cancer; genetic marker; lymphoma; cytostatic; hematological disease;
KW      immune disorder; neoplasm; colon tumor; gastrointestinal disease;
KW      adenocarcinoma; carcinoma; screening; diagnostic; prognostic;
KW      pharmaceutical; ds; gene.
XX
OS      Homo sapiens.
XX
PN      WO2005064009-A1.
PD      14-JUL-2005.
XX
PF      23-DEC-2004; 2004WO-DK000914.
XX
PR      27-DEC-2003; 2003DK-00001940.
PR      24-JAN-2004; 2004DK-0000096.
PR      07-APR-2004; 2004DK-0000586.
PR      26-NOV-2004; 2004DK-00001843.
XX
PA      (AROS-) AROS APPLIED BIOTECHNOLOGY APS.
PI      Orntoft TF, Jensen JL, Kruboffer M, Laiho P, Aaltonen LA;
PT      WPI; 2005-506661/51.
XX
PS      Claim 9; SEQ ID NO 9; 338bp; English.
XX
CC      The invention relates to a method of classifying the cancer in an
CC      individual having contracted cancer, which comprises determining the
CC      microsatellite status of the tumor. The microsatellite status or the
CC      hereditary or sporadic nature of the cancer is determined by a method
CC      comprising: (a) obtaining a sample from the individual having contracted
CC      cancer, the sample comprising gene expression products, the presence
CC      and/or amount of which forms a pattern that is indicative of the
CC      microsatellite status or the hereditary or sporadic nature of the cancer;
CC      (b) determining the presence and/or amount of the gene expression
CC      products forming the pattern, and obtaining an indication of the
CC      microsatellite status or the hereditary or sporadic nature of the cancer
CC      in the individual based on step (b). At least one or two of the gene
CC      expression products forming a pattern used to determine the
CC      microsatellite status or the hereditary or sporadic nature of the cancer
CC      are selected individually from any of the 134 genes comprising 367-7850
CC      bp (AEB22706 to AEB22839), given in the specification. Also included are:
CC      a method for treating an individual; a method for reducing malignancy of
CC      a cell; a method for producing antibodies against an expression product
CC      of a cell from a biological tissue; pharmaceutical composition, for
CC      treating a classified cancer, comprising at least one antibody, at least
CC      one polypeptide, or at least one nucleic acid and/or probe; and an assay,
CC      for classification of cancer in an individual having contracted cancer,
CC      comprising at least one marker capable of determining the microsatellite
CC      status in a sample and at least one marker in a sample determining the
CC      prognostic marker, where the microsatellite status and the prognostic
CC      marker is determined simultaneously or sequentially. The method above is
CC      useful for producing an assay for classifying cancer in animal tissue.
CC      The peptide, gene, or probe is useful for preparing a pharmaceutical
CC      composition for the treatment of a cancer in animal tissue. The method is

```

CC useful for classifying the cancer in an individual having contracted
CC cancer, i.e. colon cancer, e.g. an adenocarcinoma, a carcinoma, a
CC teratoma, a sarcoma, and/or a lymphoma. The present sequence represents a
CC colon cancer-associated gene used in the method of the invention.

XX Sequence 5028 BP; 1288 A; 1387 C; 1276 G; 1077 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	5028
Score:	5463.00	Matches:	1045
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
	14	Gaps:	0

US-09-555-342B-2 (1-1045) X AEB22714 (1-5028)

```
QY      1 MetGlyGluIleGluGlnAArgProThrProGlySerArgLeuGluValAProGluAAsnSer 20
Db      337 ATGGAGAAATATAGACAGAGGCGGACCCAGGATCAGACTGGGGGCGCCGAAATTCG 396
QY      21 GlyIleSerThrLeuGluAArgGlyGluValProProProThrProSerGlyValLeuVal 40
Db      397 GGGATCGATGACTTGGACGTCGACAGAACCGCCCAACCTTCAGGAAATCTGTG 456
QY      41 SerIleValIleGlnMetLeuAAspSerThrGlnGluValAPheGluValAProGlnAArgAla 60
Db      457 TCATCATAAATTCAGATGCTGGATGACACCCAGAGAGCATTTGAAGTTCCACAAAGAGCT 516
QY      61 ProGlyValValLeuLeuAAspAlaValCysAAsnHisLeuLeuLeuValGluGlyAAspTyr 80
Db      517 CTGGGAGAGGCTGCTGCGATGCGATGTGCATCAACCACTCAACCTCGGAGAGGCTGACTAT 576
QY      81 PheGluValLeuGluPheProAAspHisAlaValSerIleThrValAlaLeuAAspLeuLeuValPro 100
Db      577 TTGGGCTCGAGTTCTGTGATCACAAATAAGATCAGGTGTGGATCTTCCTAAACCC 636
QY      101 IleValIleGlnIleAArgProCysHisAlaValValIlySPheValIlySPhePhe 120
Db      637 ATTGTGAACAGATTAGAACGCCAAAGCACGTTGTTGTTAAGTTGTGTGAATTTCTTT 696
QY      121 ProProAAspHisThrGlnLeuGlnGluValLeuThrArgTyrIleuPheAlaLeuGlnVal 140
Db      697 CCCCCTGACACACACAACTCCAAAGAACTCACAAAGTACCTGTTCGCGCTGCAGGTG 756
QY      141 LysGluAAspLeuAlaGlnGlyAArgLeuThrCysAAsnSerThrSerAlaAlaLeuLeuIle 160
Db      757 AAGCAGGACTTGGCTCAAGGAGGTTGACGTGTATATACACACAGCGCACTCTCTTGATT 816
QY      161 SerHisIleValGlnSerGluIleGlyAAspPheAAspGluAlaLeuAAspAArgIleu 180
Db      817 TCACACTTGTGCATCTGAGATTGGGGAATTTGATGATAGCCTTGCACAGAGCACTTGA 876
QY      181 AlaIlyAAsnIlySerTyrIleProGlnGlnAAspAlaLeuGluAAspIlyIleValGluPheHis 200
Db      877 GCAAAAATAATTAATCATACTCGACAAAGACGCACTTACAGACAAATAACGGGAATTTTCAC 936
QY      201 HisAAsnHisIleGlyGlnThrProAgluSerAAspPheGlnLeuLeuGluIleAlaArg 220
Db      937 CATTAACCATTTGGACAAACACACAGCAATTCAGATTTCCAGCTCTTAGAGATTTGCCGT 996
QY      221 ArgLeuGluMetTyrGlyIleAArgLeuHisProAlaIlyAAspAArgIleuGlyThrIlyIle 240
Db      997 CGGCTGAGAGATGTATGGAATCCGGTTGCACCCGGCCAAAGACAGGAGGACAGAGATC 1056
QY      241 AsnLeuAlaValAlaAAsnThrGlyIleLeuValAPheGlnGlyPheThrIlyIleAAsnAla 260
Db      1057 AATCTGGCCCTGTGGCAACACGGGAATTCGTGTGTTTCAAGGATTCATTAAGATCAATGCCC 1116
QY      261 PheAAsnThrAlaIlyValAArgIlyLeuSerPheIlyAAspIlyValAArgPheLeuIleValLeu 280
Db      1117 TTCAACTGGGCGCAAGGTGGAGGCTGAGCTTCAAGAGGAAGCGCTTTCATCAACAGCTC 1176
```

```
QY      281 ArgProAAspAlaAAsnSerAlaTyrGlnAAspThrLeuGluPheLeuMetAlaSerAArgAAsp 300
Db      1177 CGGCGAGATGCCAATATGTCGTATCCAGAGTACCTTGGAATTCCTGATGGCCAGTCCGGAT 1236
QY      301 PheCysLeuSerPheThrIlyValIleCysValGluHisHisAlaPhePheAArgLeuPheGlu 320
Db      1237 TTCTGCAGTCTCTTCGAAATACTGTGTGAACATCATGCTCTTTAGACTTTTGAAT 1296
QY      321 GluProIlyProIlyProIlyProValAlaPheSerAArgIlySerSerPheAArgPheSer 340
Db      1297 GAGCCAAACCAAAAGCCCAAGCCCGCTCTTATGCCGGGGGTCACTATTCGGTTCACT 1356
QY      341 GlyAArgThrGlnIlyGlnValLeuAAspTyrValIlyGluGlyGlyHisIlyValGln 360
Db      1357 GCTCGGACTCAGAAAGCAGGTTCTCGACTATGTATAAGAAAGAGGACATTAAGAAAGTGCAG 1416
QY      361 PheGluAArgIlyHisSerIlyHisSerIleAArgSerLeuAlaSerGlnProThrGln 380
Db      1417 TTGAAAGAGAGCAGCAGAGATTCATCTATCCGAGGCTGTCTCACAGCTTACAGAA 1476
QY      381 LeuAAsnSerGluValLeuGluGlnSerGlnIleThrSerLeuThrPheGlyGluGly 400
Db      1477 CTGAATTCGGAAGTCTGGAGCAGTCTCAGCAGAGACCAAGCTTACATTTGAGAAAGGT 1536
QY      401 AlaGluSerProGlyGlyGlnSerCysAArgAArgIlyIlyGluProIlyValSerAlaGly 420
Db      1537 GCCGAATCTCCAGAGGGGCGCAGAGCTCGCGCAGAAAGAAAGAAAGGTTTCCGCGGG 1596
QY      421 GluProGlySerHisAProSerProAlaAProAArgAArgSerProAlaGlyAAsnIlyVal 440
Db      1597 GAGCCGGGGTCTGACCCGAGCCCTGCGCGAGGAGAAAGCCCGGGGTAAACAAAGCAGGG 1656
QY      441 AAspGlyAlaAlaSerAlaProThrGlnGluGluGluValValIlyAAspAArgThrGln 460
Db      1657 GACGAGAGCCGCTCGGCGCCACAGAGAAAGAGAGAGAGTCTTAAAGATAGCAACCG 1716
QY      461 GlnSerIlyProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeu 480
Db      1717 CAGAGTAAACCTCAGCCCGCCGACCAACACAGGCTCCCTGACTGGCAGTCCCTCACCTT 1776
QY      481 SerGluLeuSerValAAsnSerGlnGlyValAlaAlaProAlaAAsnValThrLeuSerPro 500
Db      1777 TCCGAGCTGTGTGAACCTCGCAGGGGAGAGGCGCCCTGCAACGTCGATGTCTCCC 1836
QY      501 AsnLeuSerProAAspThrIlyGlnAlaSerProLeuIleSerProLeuLeuAAsnAAspGln 520
Db      1837 AACCTAGGCCCGCAGCACCAAGACAGGCTCTCTCTTGATCAGCCGCTGCTGAATGACAG 1896
QY      521 AlaCysProAArgThrAAspGluAAspGluGlyAArgAArgIlyAArgPheProThrAAspIly 540
Db      1897 GCCTGCCCGCCGACGACATGAGATGAGGGCCGAGGAAAGATTTCCCAACTGATTA 1956
QY      541 AlaTyrPheIleAlaIlySerGluValSerThrThrGlnAArgThrTyrIleuIlyAAspLeuGlu 560
Db      1957 GCGTACTTATAGCTTAAGAAAGTGTACACAGCAACATATGTGAAGATCTCGAA 2016
QY      561 ValIleThrSerThrPheGlnSerThrValSerIlyGluAAspAlaMetProGluAlaLeu 580
Db      2017 GTTATCATCTTCGTGTTTCAAGACACAGTGAAGAAAGAGCAGCCTGCGGAGACACTG 2076
QY      581 LysSerLeuIlePheProAAspPheGluProLeuHisIlyPheHisThrAAspPheLeuIly 600
Db      2077 AAAAGTCTCATATATCCCGAATTTTGAACCTTTGCAACAAATTTTCATATTTTTCAG 2136
QY      601 GluIleGluGluAArgLeuAlaLeuThrProGluIlyAArgSerAAspAlaGlnIleAArgAAspTyr 620
Db      2137 GAAATTAAGACACGACTTCCCTGTGGAAAGCGGCTCAAAATGCCAAATTCAGAGATTA 2196
QY      621 GlnAArgIleGlyAAspValMetLeuIlyAAsnIleGlnIlyMetCysIlyHisIlyValAlaHis 640
Db      2197 CAAGAAATCGCGCATCTATCTGAAAGAACTTCAAGGAGATGAAGCACTCGGCGGCTCAC 2256
QY      641 LeuThrIlyHisSerGluAlaLeuGluAlaLeuGluAAsnGlyIleIlySerSerAArgAArg 660
```

```

Db      2257 CTGTGGAAAGCAGCAGGAGCCCTTGAGAGCCCTGAGAAATGAGAAATCAAGAGCTCCCGGGG 2316
Qy      661  LeuGluValenPhCyValArgAspPheGluLeuGlnValCyValTyrLeuProLeuAsnThr 680
Db      2217 CTGGAGAACTTCTGACAGACTTGAGAGCTGCAAGAGGTGTACTACCTGACCCCTCAACACC 2376
Qy      681  PheLeuLeuAArgProLeuHLeuLeuMetHLeuTyrTyrValLeuGlnValArgLeuCyS 700
Db      2377 TTCCTCCGGCGGCACATGCACCGGCTCATGCACTACAGCAGGCTCTGGAGCGGCTGGC 2436
Qy      701  LysHISHisProPseRHisAlaAspPheArgAspCyValArgAlaAlaLeuAlaGluLe 720
Db      2437 AAACACACACCCCGGAGCCAGCCGACCTTCAAGGACTCCAGCGGCTTTGGCAGAGATC 2496
Qy      721  ThrGluMetValAlaGlnLeuHISGlyThrMetIleLysMetGluAsnPhenGlnLysLeu 740
Db      2497 ACGGAGATGTGGCAGAGCTCCACGGTACGATGATCAAGATGAGAAATTTTCCAGAGCTG 2556
Qy      741  HisGluLeuLysValAspLeuIleGlyTyrLeuAsnLeuValAlaProGlyArgGluPhe 760
Db      2557 CACGACATCAAGAAAGATTGATGGCAATTGACAATCTTGTTCCGGAAAGGAGATTCC 2616
Qy      761  IleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnArgMetPhePhe 780
Db      2617 ATCCGCTCGGCGACCTCAGCAAGCTTCGGGGAAAGGGGCTCCAGCAGCGCATGTTCTTC 2676
Qy      781  LeuPheAsnAspValLeuLeuTyrThrSerArgLysLeuThrAlaSerAsnGlnPheLys 800
Db      2677 CTGTTCAACGACGCTCTGCTATACAGAGCCGGGGGCTGACGGCTCCAAATCAAGTTTAAA 2736
Qy      801  ValHisGlyGlnLeuProLeuTyrGlyMetThrIleGluGlnSerGluAspGluTTPGly 820
Db      2737 GTCCACGGGCGAGCTCCGCTCTATGCGATGACATGAGAGAGCAGAAAGAGAGTGGGGG 2796
Qy      821  ValProHisCyValThrTrpLeuArgGlyGlnArgInserIleIleValAlaSerSer 840
Db      2797 GTGCCCCATCTCCGACCTCCGGGGCCAGCGGCATCATCATCTGGCCGCAATGTTCT 2856
Qy      841  ArgSerGlnMetGluTyrTrpValGluAspIleGlnMetAlaIleAspLeuAlaGluLys 860
Db      2857 CGGTCCGAGATGAGAGATGGGTGAGAGACATCCAGATGGCCATTGACTGGCGAGAGAG 2916
Qy      861  SerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLysSerProAsp 880
Db      2917 AGCAGCAGCCCCCGCTGAGATCTCTGGCGAGCCCCCTGACCAACAAATCCCTCGAT 2976
Qy      881  GluAlaThrAlaAlaAspGlnGlnSerGluAspAspLeuSerAlaSerArgThrSerLeu 900
Db      2977 GAGGCCACCGCGGCTGACAGAGATCAAGAGATGACCTGAGCGGCTCCGCGACATCGCTG 3036
Qy      901  GluArgGlnAlaProHisArgGlyAsnThrMetValHisValCyValTrpHisArgAsnThr 920
Db      3037 GAGGCGCAGGCGCCGCAACCGCGCAACCAATGGTGACGTTGCTGGCACCGCAACACC 3096
Qy      921  SerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerGlyAsnLeuLeu 940
Db      3097 AGCCTCTCCATGAGGACTTCAGCATCGCATGAGATCAAGATCAAGTGTCTGGAAACCTGTG 3156
Qy      941  ArgLysPheLysAsnSerAsnGlyTrpGlnLysLeuTrpValIlePheHisAsnPhCyS 960
Db      3157 AGGAAATTCANAAACAGCAGCGGTGGCAGAAAGCTGTGGGTGCTTCACAAACTTCTGC 3216
Qy      961  LeuPhePheTyrLysSerHisGlnAspAsnHisProLeuAlaSerLeuProLeuGly 980
Db      3217 CTGTTCTTCTCAATCAACACCAAGACAAATCATCCCTTGGCCACCTGCTCTCGGCG 3276
Qy      981  TyrSerLeuThrIleProSerGlnSerGluAsnIleGlnLysAspTyrValPheLysLeu 1000
Db      3277 TACTCGCTCACCAATCCCTCTGAGCTCGAAGAACATCCAGAAAGACTGAGTTCAGAGCTG 3336
Qy      1001 HisPheLysSerHisValTyrTyrPheArgAlaGlnSerGluTyrTrpPheGluArgTrp 1020

```

```

Db      3337 CACTTCAAGTCCACGCTACTTACTTCAAGGCGGAAAGCGAGTACAGCTTCAAGAGGTGG 3396
Qy      1021 MetGluValIleArgSerAlaThrSerSerAlaSerArgProHisValLeuSerHisLys 1040
Db      3397 ATGGAAGTGAATCCGACATGGCCACCACTGCTGCGGACCCGACGTTGAGTGAACAAA 3456
Qy      1041 GluSerLeuValTyr 1045
Db      3457 GAGTCTCTGTGTAAT 3471

RESULT 4
AAA08582
ID   AAA08582 standard; DNA; 4687 BP.
AC   AAA08582;
XX
XX   19-JUL-2000 (first entry)
DE   Human cytoskeleton associated protein 2 (CYSKP-2) coding sequence.
XX
XX   Cytoskeleton associated protein; CYSKP-2; cancer; proliferative;
XX   autoimmunity; inflammatory; vesicle trafficking; neurological;
XX   cardiovascular; cell motility; reproductive; muscle disorder; ss.
XX   Homo sapiens.
XX
XX   Key Location/Qualifiers
XX   CDS 90..3227
XX   /*tag= a
XX   /product= "CYSKP-2"
XX
XX   W0200017355-A2.
XX
XX   30-MAR-2000.
XX
XX   17-SEP-1999; 99MO-US021565.
XX
XX   18-SEP-1998; 98US-0172226P.
XX   27-APR-1999; 99US-0131321P.
XX
XX   (INCY-) INCYTE PHARM INC.
XX
XX   Lal P. Tang YT, Yue H, Hillman JL, Bandman O, Corley NC;
XX   Guegler KJ, Patterson C, Azimzai Y, Baughn MR;
XX   WPI; 2000-283582/24.
XX   P-PSDB; AAY91947.
XX
XX   Human cytoskeleton associated proteins, used to treat cell proliferative,
XX   autoimmune/inflammatory, vesicle trafficking, neurological, cell
XX   motility, reproductive and muscle disorders.
XX
XX   Claim 9; Page 101-102; 113pp; English.
XX
XX   AAA08581-96 encode human cytoskeleton associated proteins 1 to 16 (CYSKP-
XX   1 to CYSKP-16) respectively. The sequences can be used to treat and
XX   diagnose cancer and cell proliferative, autoimmune/inflammatory, vesicle
XX   trafficking, neurological, cardiovascular, cell motility, reproductive
XX   and muscle disorders. Pharmaceutical compositions containing CYSKP-2 to
XX   CYSKP-16 can be used to treat or prevent disorders associated with
XX   decreased expression or activity of CYSKP (claimed), for example,
XX   atherosclerosis, cirrhosis, hepatitis, myelofibrosis, psoriasis, cancers,
XX   autoimmune/antimflammatory disorders such as allergies, anemia, asthma,
XX   acquired immunodeficiency syndrome (AIDS), Crohn's disease, diabetes
XX   mellitus, Goodpasture's syndrome, osteoarthritis, scleroderma and trauma.
XX   CYSKP antagonists can be used to treat or prevent a disorder associated
XX   with increased expression or activity of CYSKP (claimed)
XX
XX   Sequence 4687 BP; 1179 A; 1261 C; 1196 G; 1051 T; 0 U; 0 Other;
XX
Alignment Scores: 0 Length: 4687
Pred. No.: 5459.00 Matches: 1044
Score:

```

Percent Similarity: 99.9% Conservative: 0
Best Local Similarity: 99.9% Mismatches: 1
Query Match: 99.9% Indels: 0
DB: 3 Gaps: 0
US-09-555-342B-2 (1-1045) x AA08582 (1-4687)

QY 1 MetGlyGluIleGluGlnAArgProThrProGlySerArgLeuGlyAlaProGluAAsnSer 20
Db 90 ATGGAGAAATTAAGCAGAGGCCGACCCAGATACAGACTGGGGGGCCCCGGAATAATTCG 149
QY 21 GlyIleSerThrLeuGluArgGlyGlnLysProProThrProSerGlyIlySLeuVal 40
Db 150 GGATCAGTACCTTGGAACTGGACAGAAAGCCGCCCAACACTTCAGAGAAATCTCGTG 209
QY 41 SerIleValIleGlnMetLeuAAspAAspThrGlnGluAlaPheGlnValProGlnArgAla 60
Db 210 TCATCAAAATCCAGATGCTGAGATGACACCCAGAGGCAATTTGAAGTTCCAAAGAGCT 269
QY 61 ProGlyLysValIleuLeuAAspAlaValCysAAsnHisIleuAAsnLeuValGluGlyAAspTyr 80
Db 270 CTGGGAGAGGTGCTGCTGGATGCAAGTTGCAACCACTCACTCGTGGAAAGGTGACTAT 329
QY 81 PheGlyLeuGlnLysPheProAAspHisIlySLeuValIleThrValIlePheAAspLeuLysPro 100
Db 330 TTGGGCTCGAGTTTCTGTGATCACAAAAGATCACGGGTGGCTGGATCTCTAAACC 389
QY 101 ILeValIleGlnIleAAspArgProLysHisIValIValLysPheValIlyLysPhePhe 120
Db 390 ATGTGTAAACAGATTAAAGGACCAAGACGATGTGTGTAAAGTTGTGTGAATAATCTTT 449
QY 121 ProProAAspHisThrGlnLeuGlnGluGlnLysThrArgTyrLeuPheAlaLeuGlnVal 140
Db 450 CCGGCTGACCAACACAACTCCAAGAGAACTCAAGAGTACCTGTTCCGGCTGCAGGTTG 509
QY 141 LysGlnAAspLeuAlaGlnIlyArgLeuThrCysAAsnAAspThrSerAlaAlaLeuLeuIle 160
Db 510 AACGAGACATTTGGCTCAAGGACAGGTTGACGTTAATACACAGACGACGCTCTTGATT 569
QY 161 SerHisIleValGlnSerGluIleGlyAAspPheAAspGluAlaLeuAAspArgGluHisIleu 180
Db 570 TCAACATTTGTGCAATGTGAATGGGAAATTTGATGAAGCTTTGGACAGAGACCTTA 629
QY 181 AlaLysAAsnLysTyrIleProGlnGlnAAspAlaLeuGlnAAspLysIleValGluPheHis 200
Db 630 GCAGAAAATTAATACATCTCAAGCAAGACGCACTAAGACAAATCGTGAATTTCAAC 689
QY 201 HisAAsnHisIleGlyGlnThrProAlaGluSerAAspPheGlnLeuLeuGluIleAlaArg 220
Db 690 CATTAACCACTTGGACAAACACAGCAAGCAATTCAGATTTCCAGGCTCCPAGAGATTCGCCGT 749
QY 221 ArgLeuGlnMetTyrGlyIleArgLeuHisProAlaLysAAspArgGluGlyThrLysIle 240
Db 750 CGGCTGAGAGATGTAATGGAATCCGGTTGACCCGGCCAAAGACAGGAGGACAGAGATC 809
QY 241 AsnLeuAlaValAlaAAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAAsnAla 260
Db 810 AATCTGGCCGTTGGCAACAGGAAATTCAGTGTTCAGGGTTTCACTAAGATCAATGCC 869
QY 261 PheAAsnThrAlaLysValIleArgLysLeuSerPheLysAArgLysArgPheLeuIleLysLeu 280
Db 870 TTCAACCTGGGACCAAGGTCGGAAGCTGAGCTTCAAGGAGAAAGCGCTTTTTCATCAAGCTC 929
QY 281 ArgProAAspAlaAAsnSerAlaTyrGlnAAspThrLeuGlnLysPheMetAlaSerArgAsp 300
Db 930 CGGCGAATGCAATTAATGCTGATCACAGATACCTTGAATTCCTGATGCGCACTCGGAGAT 969
QY 301 PheCysLysSerPheThrLysIleCysValGluHisHisAlaPhePheArgLeuPheGlu 320
Db 990 TTCTGCAAGTCTTCTGAAATCTGTGTGAACATCATGCTCTTTAAGCTTTTGA 1049
QY 321 GluProLysProLysProLysProValLeuPheSerArgLysSerPheArgPheSer 340

Db 1050 GAGCCCAAAACAAAGCCCAAGCCGCTCTTTAGCCGGGGGTCAATTCATTCGTTCAAGT 1109
QY 341 GlyArgThrGlnLysGlnValLeuAAspTyrValIlyGlyGlyHisIlySLeuValGln 360
Db 1110 GGTTCGACTCAGAGAGAGGTTCTCGACTATGTTTAAAGAGAGACATTAAGAGGTCCAG 1169
QY 361 PheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
Db 1170 TTGAAGAGAGACACAGAAAGATTCATTCATCCGAGCCTTGCTTCAAGCTTAAGAA 1229
QY 381 LeuAAsnSerGluValLeuGlnSerGlnGlnIleSerThrSerLeuThrPheGlyGlyGly 400
Db 1230 CTGAATTCGAAGTGTGAGAGAGTCTCAGAGAGACCAAGCTTTCATTTGAGAAAGGT 1289
QY 401 AlaGluSerProGlyGlyGlnSerCysArgArgLysGlyLysProLysValIleSerAlaGly 420
Db 1290 GCCGATCTTCAGGGGGCCAGAGCTCCGGCGAGAGAAAGAACAGAGTTTCCGCCGG 1349
QY 421 GluProGlySerHisAAspSerProAlaProArgArgSerProAlaLysAsnLysGlnAla 440
Db 1350 GAGCCGGGGTCCGACCCGAGCCCTGGCCGAGAGAGAACCCCGGGGTAAACAGAGCGC 1409
QY 441 AAspGlyAlaAlaSerAlaAAspThrGlnGlnGlnGlnValValLysAAspArgThrGln 460
Db 1410 GACGAGCCGCTCGGCGGCCACGAGAGAGAGAGAGAGTGTAAAGATTAAGACCAAG 1469
QY 461 GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeu 480
Db 1470 CAGAGTAACTCAGGCCCGCCGACGACCAACAGAGCTCTTGACTGAGCTTCACTT 1529
QY 481 SerGlnLeuSerValAAsnSerGlnGlyIlyValAlaProAlaAAsnValIleLeuSerPro 500
Db 1530 TCCGAGCTGCTGTGAATCGACAGGGGGAGTGGCCCTGCGCAACGTGACTGTCTGCC 1589
QY 501 AAsnLeuSerProAAspThrLysGlnAlaSerProLeuIleSerProLeuAAsnAAspGln 520
Db 1590 AACCTAGGCCGACACCAAGAGGCTCTCTCTTATACAGCCGCTGTAAATGACCAAG 1649
QY 521 AlaCysProArgThrAAspArgGluAAspGluGlyArgArgLysAAspPheProThrAAspLys 540
Db 1650 GCTGCCCCCGAGCGAGCATAGAGATGAGGGCCGAGAGAAAGATTCCCACTGATTA 1709
QY 541 AlaTyrPheIleAlaLysGlnValSerThrThrGlnArgThrTyrLeuLysAAspLeuGlu 560
Db 1710 GCCTACTTATAGCTAAGAGAGTGTCTACACGAGACGAAACATATGTGAAGGATTCGAA 1769
QY 561 ValIleThrSerThrPheGlnSerThrValSerLysGlnAAspAlaMetProGluAlaLeu 580
Db 1770 GTTATCACTTCGTGTTTCAAGACAGAGTGAAGAAAGAGACGCAATGCCGGAAGCACTG 1829
QY 581 LysSerLeuIlePheProAAspPheGluProLeuHisIlySLeuPheHisThrAAspPheLys 600
Db 1830 AAAAGTCTCATATTCGGAATTTTGAACCTTTGCAAAATTTCAATTAATTTTCCAAAG 1889
QY 601 GluIleGlnGlnArgLeuAlaLeuTyrGlnGlyArgSerAAsnAlaGlnIleArgAAspTyr 620
Db 1890 GAAATTGACCAACGACTTCCCTGCTGGGAAAGCCGCTCAAAATGCCCCAAATCAAGATTAAC 1949
QY 621 GlnArgIleGlyAAspValMetLeuLysAAsnIleGlnIleMetLysHisIleuAlaHis 640
Db 1950 CAAAGATCGCCCATATCATGTGAAAGACATTCAGAGGCACTTGAAGCACTCGGGGCTCAC 2009
QY 641 LeuTyrLysHisSerGlnAlaLeuGlnAlaLeuGluAAsnGlyIleLysSerSerArgArg 660
Db 2010 CTGTGAGACACAGCGAGGCTTTGAGGGCCCTGGAGAAATGGAATCAAGAGCTCCGGCGG 2069
QY 661 LeuGlnAAspPheCysArgAAspPheGluLeuGlnLysValCysTyrLeuProLeuAAsnThr 680
Db 2070 CTGAGAAACTTTCGACAGACTTTGAGCTGCAAGAGGTGTGTAACTTACCGCTCAACACC 2129
QY 681 PheLeuLeuArgProLeuHisIleArgLeuMetHisTyrLysGlnValLeuGluArgLeuCys 700
Db 2130 TTCTCTCGGGCACTGACCGGCTCATGCACTAACAAGAGTCTTGAAGCGGCTGTGC 2189

```

QY 701 LysHisAspProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIle 720
DB 2190 AATACACACCCCGGAGCCAGCCGACTTCAGAGCTGCGAGCGCTTGTGGCAGAGTTC 2249
QY 721 ThrGluMetValAlaGluLeuHisGlyThrMetIleLysMetGluAsnPheGlnLysLeu 740
DB 2250 ACCGAGATGGTGGCAGAGCTCCACGGTACGATGATCAAGATGGAGAAATTTCCAGAGAGTGG 2309
QY 741 HisGluLeuLysLysAspLeuIleGlyIleAsnAsnLeuValValProGluArgGluPhe 760
DB 2210 CACCAACTCAGAGAAAGATTGATGGCATGACATCTTGGTGTCCGGGAAAGGAGATTC 2369
QY 761 IleArgGluGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnAsnMetPhePhe 780
DB 2370 ATCCGCTGGGCGACCTCCAGAGAGCTCTGGGAGAGGGGCTCCAGCAGCGCATGTCTTC 2429
QY 781 LeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsnGlnPheLys 800
DB 2430 CTGTTCAACGACGTCCTGCTATACACGAGCGGGGGCTGAGCGGCTCCAATCAGTTTAA 2489
QY 801 ValHisGlyGluLeuProLeuTyrGlyMetThrIleGluGluSerGluAspGluTyrGly 820
DB 2490 GTCCACCGGCGAGCTCCGCTCTATGGCATGAGATTGGAGAGACGAAAGAGAGTGGGGG 2549
QY 821 ValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAlaSerSer 840
DB 2550 GTGCGCCACCTGCTGAGCCCTCGGGGGCCAGCGGAGTCCATCATCGTGGCCGCAATCT 2609
QY 841 ArgSerGluMetGluLysTyrPheValGluAspIleGlnMetAlaIleAspLeuAlaGluLys 860
DB 2610 CGGTCCGAGATGGAGATGGGTGAGAGACATCCAGATGGCCATTGACCTGGCGAGAGAG 2669
QY 861 SerSerSerProAlaProGluPheLeuAlaSerSerProAspAsnLysSerProAsp 880
DB 2670 AGCAGCAGCCCCCGCTGAGATCTCTGGCCAGCAGCCCCCTGCAACAATCTCCCTGAT 2729
QY 881 GluAlaThrAlaAlaAspGlnGluSerGluAspAspLeuSerAlaSerArgThrSerLeu 900
DB 2730 GAGGCCACCGGGGCTGACAGAGATCAGAGATACCTGAGCGGCTCGCCGACATCCCTG 2789
QY 901 GluArgGlnAlaProHisArgGlyAsnThrMetValHisValCysTyrPheArgAsnThr 920
DB 2790 GAGGCGCAGCGCCCGCAGCGCGGACACACATGGTGACGTTGCTGGCAGCGGACACC 2849
QY 921 SerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerGlyAsnLeuLeu 940
DB 2850 AGCGTCTCATGTGGACTTCAGCATCGCAGTGGAGATCAGTTGTGTGGAACCTGCTG 2909
QY 941 ArgLysPheLysAsnSerHisGlyTyrGlnLysLeuTyrPheValValPheThrAsnPheCys 960
DB 2910 AGGAAATTCACAAACACACACAGGGGTGGAGAAAGTGGGTGGTTCACAAACTTCTGC 2969
QY 961 LeuPhePheTyrLysSerHisGlnAspAsnHisProLeuAlaSerLeuProLeuLeuGly 980
DB 2970 CTGTTCTTTCACAAATCACACAGACATATCCCTTCGCGACCTGCTCTCTGGGC 3029
QY 981 TyrSerLeuThrIleProSerGluSerGluAsnIleGlnLysAspTyrValPheLysLeu 1000
DB 3030 TACTCGCTCACATCCCTCGATCGAGTCCGAGAACATCCGAAAGACTAGTTCACAGCTG 3089
QY 1001 HisPheLysSerHisValTyrTyrPheArgAlaGluSerGluTyrThrPheGluLysTyr 1020
DB 3090 CACTTCAGATCCCGCTCTACTTACTTCAAGGGCGGAAAGCGATACAGTTCGAAAGGTGG 3149
QY 1021 MetGluValIleAspSerAlaThrSerSerAlaSerArgProHisValIleLysSerHisLys 1040
DB 3150 ATGGAATGATCCCGCAGTCCACAGCTCTGCTCTCGGACCCCACTGTGAGTCACAAA 3209
QY 1041 GluSerLeuValTyr 1045
DB 3210 GAGTCTCTGTGTAT 3224

```

```

RESULT 5
ID AAS64760 standard; cDNA; 3187 BP.
AC AAS64760;
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #564.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
DR WP1; 2001-639362/73.
PS P-PSDB; ABG00573.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 564; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3187 BP; 810 A; 874 C; 876 G; 627 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0	3187	1042	0	3	1	0
Score:	5431.00					
Percent Similarity:	99.6%					
Best Local Similarity:	99.6%					
Query Match:	99.4%					

US-09-555-342B-2 (1-1045) x AAS64760 (1-3187)

1 MetGlyGluIleGluGlnArgProThrProGlySerArgLeuGluAlaProGluAsnSer 20

```

|||||
Db 49 ATGGAGAAATAGAGCGAGCGACCCCGAGATCACGACTGGGGGCCCCGGAATTCG 108
Qy 21 G|Y|I|S|e|r|h|r|e|u|g|u|a|r|g|l|y|g|l|y|a|P|r|o|P|r|o|h|r|P|r|o|s|e|r|g|l|y|l|e|u|V|a| 40
Db 109 GGGATTCGTAACCTTGACAGCTGGACAGAGCCGCCCAACACCTTCAGGAAACCTCGTG 168
Qy 41 S|e|r|l|e|y|s|i|e|g|i|m|e|t|e|u|a|a|p|r|t|h|r|g|i|n|u|a|l|e|h|e|g|u|a|l|P|r|o|g|i|n|a|r|g|a| 60
Db 169 TCCATCAAAATCCAGATGCTGGATGACACCCAGAGGACATTTGAAAGTTCCACAAAGGCT 228
Qy 61 P|r|o|g|i|l|y|s|a|l|l|e|u|e|u|a|a|v|a|l|C|y|a|a|h|h|i|s|l|e|u|a|n|l|e|u|a|l|g|i|u|g|y|a|s|r|t|y|r| 80
Db 229 CCGGGAAAGGTGCTGCTGAGATGACAGTTTGCAACCACTCAACCTCGGGAAAGGTGACTAT 288
Qy 81 P|h|e|g|l|y|l|e|u|g|l|u|p|h|e|r|P|r|o|a|s|h|i|s|l|y|s|l|e|h|r|v|a|l|t|r|p|h|e|u|a|p|h|e|u|l|e|u|y|P|r|o| 100
Db 289 TTTGGCTCGAGTTTCTTGATCACAAAAAGATCAGGTGGCTGGATCTCTTAAACCC 348
Qy 101 I|l|e|v|a|l|y|s|g|i|n|l|e|a|r|g|a|r|P|r|o|l|y|h|i|s|v|a|l|v|a|l|y|s|P|h|e|v|a|l|l|y|s|P|h|e| 120
Db 349 ATTGTGAACAGATTAAAGGCCAAAGCACGTTGTTGTTAAAGTTTGTGGTAATTCCTT 408
Qy 121 P|r|o|P|r|o|a|s|h|i|v|h|r|g|i|n|l|e|u|g|i|u|l|e|u|h|r|a|r|g|t|r|l|e|u|p|h|e|a|l|a|l|e|u|g|i|n|v|a|l| 140
Db 409 CCGCCTGACCAACACAACTCCAAAGAACTCAAGAGTACCTGTTCCGGCTGCAAGTGC 468
Qy 141 l|y|s|g|i|a|s|r|p|h|e|u|l|a|g|i|n|g|l|y|a|r|g|l|e|u|h|r|C|y|a|a|h|h|r|t|h|r|s|e|r|a|l|a|l|a|l|e|u|l|e| 160
Db 469 AAGCAGACTTGGCTCAAGCAGGTTGACGTGTATACACAGCGACGCTCTTGATTT 528
Qy 161 S|e|r|h|i|l|e|v|a|l|g|i|n|s|e|r|g|i|u|l|e|g|y|a|s|r|P|h|e|a|p|h|u|a|l|a|l|e|u|a|r|g|a|r|g|i|u|h|i|s|l|e|u| 180
Db 529 TCACACATTTGCAATCTGAAGATTGGGATTTTGATGAAGCTTTGACAGAGCACTTA 588
Qy 181 A|l|e|l|y|a|h|e|u|l|y|r|t|i|l|e|r|o|g|i|n|g|i|n|a|p|h|a|l|e|u|g|i|u|a|s|r|l|y|s|i|l|e|v|a|l|g|i|u|p|h|i|s| 200
Db 589 GCAAAATTAATACATACCTCAGCAAGACGCACTAAGAGCAAAATCGGGAATTTTCAC 648
Qy 201 H|l|a|s|h|i|s|i|l|e|g|i|n|h|r|P|r|o|a|l|g|l|u|s|r|a|s|r|P|h|e|g|i|n|l|e|u|g|i|u|h|i|e|a|l|a|r|g| 220
Db 649 CATACACATTTGGACAAACACAGCAGATCAGATTTCCAGCTCTTACAGATTTGCCCT 708
Qy 221 A|r|g|l|e|u|g|i|u|e|r|t|r|g|i|l|e|a|r|g|l|e|u|h|i|s|r|o|a|l|y|a|s|r|a|r|g|i|u|g|i|l|y|h|r|y|i|l|e| 240
Db 709 CGGCTAGAGATGATGGAATCCGGTTCACCCGGCCAAAGACAGGAAAGCACAAGATC 768
Qy 241 A|h|l|e|u|h|a|l|a|a|h|h|r|g|i|l|e|l|e|u|v|a|l|P|h|e|g|i|n|g|i|l|y|h|r|h|r|y|i|l|e|a|h|i|a| 260
Db 769 AATCTGGCCCTTGCAACAGGGAATTTCTAGTGTTCAGGGTTTCACTTAAGATCAATGCC 828
Qy 261 P|h|e|a|s|h|r|a|l|a|y|s|v|a|l|a|r|g|l|y|s|l|e|u|s|r|P|h|e|l|y|a|r|g|l|y|s|r|P|h|e|u|i|l|e|y|l|e|u| 280
Db 829 TTTCACTGGGCCAAAGGTGGAGAGCTGAGCTTCAAGAGAAAGCCTTTTCTCAACAGCTC 888
Qy 281 A|r|g|r|P|r|o|a|s|h|i|a|a|h|s|e|r|a|l|a|r|g|i|n|a|s|r|t|h|r|l|e|u|g|i|u|h|e|u|h|e|a|l|a|s|r|a|r|g|a|r| 300
Db 889 CGGCGAGATCCCATAGTGGCTACAGAGATACCTTGGAATTTCTGATGGCCAGTCCGGAT 948
Qy 301 P|h|e|C|y|l|y|s|r|P|h|e|r|t|r|l|y|i|l|e|C|y|s|v|a|l|g|i|u|h|i|s|i|s|a|l|a|h|e|r|h|e|a|r|g|l|e|u|h|e|g|i|u| 320
Db 949 TTTGTGCAAGTCTTCTGGAAATCTGTGTGAACATCATCTCTTTTAAACTTTTGGAA 1008
Qy 321 G|l|u|P|r|o|l|y|s|r|P|r|o|l|y|s|r|P|r|o|v|a|l|l|e|h|P|h|e|s|r|a|r|g|i|l|y|s|r|P|h|e|r|P|h|e|r|P|h|e|r| 340
Db 1009 GAGCCCAACCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGGTCATATATTCGGTTCAGT 1068
Qy 341 G|l|y|a|r|g|h|r|g|i|n|l|y|s|g|i|n|v|a|l|l|e|u|a|s|r|t|r|y|r|v|a|l|y|s|g|i|u|g|i|l|y|h|i|s|l|y|l|y|s|v|a|l|g|i|n| 360
Db 1069 GGTCCGAGCTCAGAGAGGTTCTCGACTATGTTAAAGAGAGAGACATTAGAAAGGTCCAG 1128
Qy 361 P|h|e|g|i|u|a|r|g|l|y|h|i|s|e|r|l|y|i|l|e|h|i|s|e|r|i|l|e|r|g|s|e|r|i|e|u|a|l|a|s|r|g|i|n|P|r|o|h|r|g|i|u| 380

```

```

Db 1129 TTTGAAAGAAAGACAGCAAGATTCATTCCTGGAGCCTTGCTTCACAGCTACAGAA 1188
Qy 381 l|e|u|a|h|s|e|r|g|i|u|v|a|l|l|e|u|g|i|n|s|e|r|g|i|n|s|e|r|h|r|P|h|e|l|y|g|i|u|-g|i| 400
Db 1189 CTGAATTCGGAAGTCTGAGAGCATGCTCAGCAGAGACACAGCTTTCATTTGGAGAGAGG 1248
Qy 400 y|a|l|a|g|l|u|s|r|P|r|o|g|i|l|y|g|i|n|s|e|r|C|y|a|r|g|a|r|g|i|l|y|s|g|i|u|p|r|o|l|y|s|a|l|s|e|r|a|l|a|g|i| 420
Db 1249 TCCGGAATTTCCAGGGGGCCAGAGCTGCGGCGAAGAAAGAAACCAAGAGTTTCCGCCGG 1308
Qy 420 y|g|i|u|p|r|o|g|i|y|s|r|h|i|s|r|P|r|o|s|r|P|r|o|a|r|g|a|r|g|s|r|P|r|o|a|l|g|l|y|a|n|l|y|s|g|i|n|l| 440
Db 1309 GAGCCGGGGGTGCACCCAGCCCTGCGCGAAGAAAGCCCGCGGTAAACAGAGGC 1368
Qy 440 a|a|e|r|g|i|y|a|l|a|l|a|s|e|r|a|l|a|r|h|r|g|i|u|g|i|u|g|i|u|v|a|l|l|y|s|h|a|r|g|h|r|g|i| 460
Db 1369 GAGCGAGAGCCGCTCCGAGCCCAAGGAAAGAGAGAGTGTGTTAAGATAGAACCA 1428
Qy 460 n|g|i|n|s|e|r|l|y|s|r|P|r|o|g|i|n|P|r|o|g|i|n|P|r|o|s|r|h|r|g|i|y|s|r|l|e|h|r|P|h|e|r|P|h|i|s|l|e| 480
Db 1429 GCAAGATTAACCTCAAGCTCCGACGCCAAGCAGAGCTCCTGATCGCAGTCCGACCT 1488
Qy 480 u|s|r|g|i|u|e|u|s|r|v|a|l|a|h|s|e|r|g|i|n|g|i|l|y|v|a|l|a|r|P|r|o|a|a|h|v|a|l|t|h|r|l|e|u|s|r|P|r| 500
Db 1489 TTTCCAGCTGTCTGTAACTCGCAGGGGAGTGGCCCTGCGCAAGCTGACCTTGCTCC 1548
Qy 500 o|a|h|l|e|u|s|r|P|r|o|a|s|h|r|t|h|r|l|y|s|g|i|n|a|l|a|s|r|P|r|o|l|e|u|i|l|e|s|r|P|r|o|l|e|u|a|h|a|r|g|i| 520
Db 1549 CAACCTGAGCCCGCAACACAGAGGCTCTCCCTTGATCAAGCCCTGCTGAATGACCA 1608
Qy 520 n|a|l|a|C|y|P|r|o|a|r|g|h|r|a|h|a|r|p|h|u|a|r|g|i|u|g|i|l|y|a|r|g|l|y|s|a|r|g|P|h|e|r|h|r|P|h|e|r|h|y| 540
Db 1609 GGCCTGCCCCCGGACGAGATGAGGATGAGGCGCGAAGAAAGATTTCCAACTGATTA 1668
Qy 540 s|a|l|a|r|P|h|e|i|l|e|a|l|y|s|g|i|u|v|a|l|s|e|r|h|r|h|r|g|i|u|a|r|g|h|r|t|h|r|l|e|u|h|y|s|a|r|p|h|e|u|g|i| 560
Db 1669 AGGTACTTCATAGCTTAAGAAAGTGTCTACACCGAGCAACATATCTGAAGATCTCGA 1728
Qy 560 v|a|l|l|e|h|r|h|r|s|r|t|r|P|h|e|g|i|n|s|r|h|r|v|a|l|s|e|r|l|y|s|g|i|u|a|s|r|a|l|a|h|e|r|P|r|o|g|i|u|a|l|e| 580
Db 1729 AGTTATCACTCTTGCTGTTTCAAGCAGCAGTGAAGCAAGAGAGCCGATGCGGAACACT 1788
Qy 580 u|l|y|s|s|r|l|e|u|i|l|e|h|r|P|r|o|a|h|P|h|e|g|i|u|P|r|o|l|e|h|i|s|l|y|s|r|h|i|h|r|t|h|r|a|h|P|h|e|l|e|u| 600
Db 1789 GAAGATCTCATATATCCCAATTTTGAACTTTGCAAAATTTCAATTAATTTTCTCAA 1848
Qy 600 e|g|i|u|l|l|e|g|i|u|g|i|n|a|r|g|l|e|u|a|l|e|u|h|r|g|i|u|g|i|l|y|a|r|g|s|r|a|h|a|l|g|i|l|l|e|h|r|a|r|t|y| 620
Db 1849 GGAATTTGAACAGACTTGCCCTGTGGAAAGCCGCTCAAAATGCCAAATTCAGAGATTA 1908
Qy 620 r|g|i|n|a|r|g|i|l|e|g|y|a|s|r|v|a|l|l|e|u|l|y|s|a|h|i|l|e|g|i|l|y|e|h|i|s|l|e|u|h|a|l|a|h|i| 640
Db 1909 CCAAAATTCGGGAGTATGCTGTGAAGAACTTCAAGGCGATGAACACTGCGCGGCTCA 1968
Qy 640 s|l|e|u|h|r|l|y|s|h|i|s|e|r|g|i|u|a|l|e|u|g|i|u|a|h|e|u|g|i|u|a|h|e|u|g|i|l|y|l|e|y|s|r|s|e|r|a|r|g|a|r| 660
Db 1969 CCGTGTGAAGACACAGCGAGGCTTGAGAGCCCTGAGGAATGGAAATGGAATCAAGAGCTCCGGCGG 2028
Qy 660 g|l|e|u|g|i|u|a|h|P|h|e|r|C|y|a|r|g|a|r|P|h|e|g|i|u|e|u|g|i|n|l|y|s|v|a|l|C|y|t|r|l|e|u|P|r|o|l|e|h|a|h|h| 680
Db 2029 GCTGGAGAACTTTCAGAGACTTTGAGCTGAGAAAGGTGTGTACTTACCTGCTCAACAC 2088
Qy 680 r|h|e|l|e|u|h|a|r|P|r|o|l|e|h|i|s|r|g|l|e|h|u|h|e|h|i|s|r|t|r|l|y|s|g|i|n|v|a|l|l|e|u|g|i|u|a|r|g|l|e|u|C|y| 700
Db 2089 CTTTCCCTCGCGGCACTGCAAGGCTCATGCACTCAAGCAGTCTGGAAGCGGCTGTG 2148
Qy 700 s|l|y|s|h|i|s|h|i|s|r|P|r|o|P|r|o|s|r|h|i|s|a|l|a|h|P|h|e|r|g|a|r|P|r|C|y|a|r|g|a|l|a|l|e|u|h|a|g|i|u|i|l| 720
Db 2149 CAACACACACCCGCGAGCAGCGCACTTCAGGGCTCCGAGCGCTTTGGCAGAGAT 2208
Qy 720 e|h|r|g|i|u|h|e|C|v|a|l|a|g|i|l|e|h|i|s|g|i|l|y|h|r|h|e|h|i|l|y|s|e|r|g|i|u|a|h|P|h|e|g|i|n|l|y|s|e| 740
Db 2209 CAGGAGATGGTGGCAGAGCTCCAGGTACGATGATCAAGATGAAGATTTCCAGAACT 2268

```



```

Oy 740 uHiegluLeuLySLySaPleuIleGIYIleaspaNleuValIProGIYARGIuPh 760
Db 2269 GCAAGAACTCAAGAAAGATTGATTGGATTCGACAAATCTTGAGTTCGGGAGAGGAGATT 2328
Oy 760 eIlaRgluGIySerLeuSerLySLeUSerGIYLySGIyLeuGIInGIInaRgeMetPhePh 780
Db 2239 CATCGCTGGGAGAGCTCTCAGCAAGCTCTCGGGAGAGGGGCTCCAGCAGCGCATTTCTT 2388
Oy 780 eLeuPheAaAaPValIleuLeuTYrThSeArGIyLeuThraIaSerAaGIInPheLy 800
Db 2289 CCGTTCACACACGCTCGCTATACACAGCGCGGGGCTGACGCGCTCAATCGATTAA 2448
Oy 800 aValHieGIyGIInePLeuTYrGIYMeTThrIleGIuGIuSerGIuAaPGIuTrGI 820
Db 2449 AGTCACGCGGAGCTCCGCTCTATGSCATGACATTAAGAGAGAGAGAGAGAGAGAGAG 2508
Oy 820 YValProHieGIyLeuThraLeuArGIYGIInaRGIInSerIleIleValIaIaSerSe 840
Db 2509 GGTGCCCACTGCTGACCTCCGGGGCCAGCGGCACTCATCTCGCTGCGCCAGTTCC 2568
Oy 840 rARSeGIuMeGIuLyTrPValGIuAaPPIeGIInMeCAIleAaPLeuAIGIuLy 860
Db 2569 TCGGTCCGAGATGAGAAAGTGGTTCAGAGACATCCAGATGCGCATTCGCGCGAGAA 2628
Oy 860 aSerSeSerProIaPProGIuPheLeuAaSerSeSerProIaPProAaPLeuYSerProA 880
Db 2629 GAGACGACGCGCGCGCTGAGTCTTGCCAGCAGCGCGCGCTGCAACAGATCCCTCGA 2688
Oy 880 PGIuAlaThraIaIaAaPGIInGIuSerGIuAaPLeuSerAaSerAaTrSerSe 900
Db 2689 TGAAGCCACCGCGCTGACAGAGTCAAGAGATACCTGAGCGCTGCGCCACATCCCT 2748
Oy 900 uGIuARGIInaIaProHieArGIYArGIYArThMeValHieValCYrTPHISArGArTh 920
Db 2749 GAGACGCGACGCGCGCGCGCGCAACATGCTGCACTGCTGCGACCGCGCAACAC 2808
Oy 920 rSerValSerMeValaAaPheSerIleIaIaValGIuAaGIIneUSerGIYArGIuLe 940
Db 2809 CAGGCTCTCCAGTGGAGCTTCAGATGCGAGTGAAGATCACTGCTGGAACCTGCT 2868
Oy 940 uARGIyPheLySaAaSerAaGIYTrGIInLyLeuTrPValIaPheThraAaPheCY 960
Db 2869 GAGAAATTCAAAACAGCAACGGGTGCAAGCTGTGGGTGCTTCAACAACTTCTG 2928
Oy 960 eLeuPhePheTYrLySeSerHieGIInaPAsNHiePLeuAaSerLeuPLeuLeuGI 980
Db 2929 CCGTCTCTTCAAAATCACAACAGCAACATCATCCCTGCGCGCTGCTGCTCGG 2988
Oy 980 YTrSerLeuThraIePProSerGIuSerGIuAaMIleGIInYArPTrValIaPheLyLe 1000
Db 2989 CTACTCGCTCAACATCCCTCTGAGTCCGAGAAACATCCAGAAAGACTACGTTCAACCT 3048
Oy 1000 uHiePheLySerHieValTYrTYrPheArGIaGIuSerGIuTYrThPheGIuArGI 1020
Db 3049 GCATTCACAGCTCCACGCTCACTACTTACGGCGGAAAGCAAGTACAGCTTCGAAGGTG 3108
Oy 1020 pMeGIuValIleArGISeArIaThrSeSerAaSerArPProHieValIeUSerHieLy 1040
Db 3109 GATGGAAGTATCCGCGAGTGCACACAGCTGCGCTCGGAGCCCACTGTGAGCCACAA 3168
Oy 1040 eGIuSerLeuValTYr 1045
Db 3169 AGAGTCTCTGTGTAT 3184

```

RESULT 6
AD084807
ID AD084807 standard; cDNA; 3997 BP.

XX AD084807;
XX
XX 07-OCT-2004 (first entry)
XX

DE Human tumour-associated antigenic target (TAT) cDNA sequence #1621.
XX human; tumour-associated antigenic target; TAT; cytotoxic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX Homo sapiens.
XX WO2004060270-A2.
XX 22-JUL-2004.
XX 15-OCT-2003; 2003WO-US029126.
XX 18-OCT-2002; 2002US-0418988P.
XX (GETH) GENENTECH INC.
XX (WO/ID/) WO T D.
XX (ZHOU/) ZHOU Y.
XX Wu TD, Zhou Y;
XX WPI; 2004-534300/51.
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
XX Claim 1; SEQ ID NO 1621; 5504pp; English.
XX
XX The present invention describes an isolated tumour-associated antigenic
XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
XX sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
XX (c). Also described: (1) an expression vector comprising the above
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)
XX a process for producing a polypeptide; (4) an isolated polypeptide
XX comprising: (a) an amino acid sequence encoded by any of the above
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-
XX length coding region of the above nucleotide sequences; or (c) a sequence
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)
XX an isolated antibody that binds to the above polypeptide; (7) a process
XX for producing the antibody; (8) an isolated oligopeptide that binds to
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)
XX binding organic molecule that binds to the above polypeptide; (10) a
XX composition of matter comprising the above (chimeric) polypeptide,
XX antibody, oligopeptide or TAT binding organic molecule, in combination
XX with a carrier; (11) an article of manufacture comprising a container and
XX the composition of matter contained within the container; (12) methods of
XX inhibiting the growth of a cell that expresses the above protein, where
XX the growth of the cell is at least in part dependent upon a growth
XX potentiating effect of the above protein; (13) a method of
XX therapeutically treating a mammal having a cancerous tumour comprising
XX cells that express the above protein; (14) a method of determining the
XX presence of a protein in a sample suspected of containing the protein
XX described above; (15) methods of diagnosing the presence of a tumour in a
XX mammal; (16) a method for treating or preventing a cell proliferative
XX disorder associated with increased expression or activity of the above
XX protein; and (17) a method of binding an antibody, oligopeptide or
XX organic molecule to a cell that expresses the protein described above.
XX The TAT sequences have cytostatic activities, and can be used in gene
XX therapy. The composition and methods are useful for diagnosing,
XX preventing or treating cancer. The composition is also used for preparing
XX a medicament for the therapeutic treatment or diagnostic detection of a
XX cell proliferative disorder or cancer. The present sequence represents a
XX human TAT cDNA sequence from the present invention.

Alignment Scores:
Pred. No.: 1,31e-189 Length: 3997
Score: 2941.50 Matches: 580
Percent Similarity: 70.0% Conservative: 162

Best Local Similarity: 54.7% Mismatches: 261
Query Match: 53.8% Indels: 57
DB: 12 Gaps: 14
US-09-555-342b-2 (1-1045) x AD084807 (1-3997)

Qy 1 MetGlyGluIleGluGlnArg-----ProThrProGlySerArgLeuGlyAla 16
Db 118 ATGGGGAGATGAGAAAGACATACAGAGTCTCTGACAGCTGACAGGAGATGCCCTTGCGGTC 177

Qy 17 ProGluAsnSerGlyIleSerThrLeuGluArgGlyGluIleProProThrProSer 36
Db 178 CAGACCCCTGGAGGATTAGACCCCTTGAGCTGGGACACTCTCTTCCCGACATGCCAA 237

Qy 37 GlyLeuValSerIleValIleGluMetLeuAspArgThrGlnIleAlaPheGluVal 56
Db 238 GAGAGACCTGGACCTCAGAGTAAAGCTGCTGACACACCATGGAATATTGACATT 297

Qy 57 ProGlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuVal 76
Db 298 GAGCTTAATAGCATGGCCAGGTATTACTGACACAAATGGGAAGCCTTTAACTGGTA 357

Qy 77 GluGlyAspTyrPheGlyLeuGluPheProAspHisValLysValIleThrValTyrLeuAsp 96
Db 358 GAATGTGACTACTTCGGGATGAGTTTCAAAATCTGACTGACTGATTTGGCTTGAA 417

Qy 97 LeuLeuSerProIleValLysGlnIleArgArgProLysHisValValLysPheVal 116
Db 418 CTTATGAAACCATCATTTAGCCAAATACGAAAGCCAAAGATGGTTCGCTTCCGCTAGCT 477

Qy 117 ValLysPhePheProProAspHisThrGlnLeuGlnGluLeuThrArgTyrLeuPhe 136
Db 478 GTAAATATTTTTCACCTGATCTCGTCAGCTACAAAGAAATATACAAAGATCTGTTT 537

Qy 137 AlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAla 156
Db 538 GCCTTGCATCTTACGAGAGACCTGCTGGAAGAGGTTTGACTGTGTGACACCCAGCG 597

Qy 157 AlaLeuLeuLeuSerHisValIleGlnSerGluIleGlyAspPheAspGluAlaLeuAsp 176
Db 598 GCCCTTCTCAGTCCCTCCCTCTGCACTCGAAATAGAGATTCAGTAAACCTTGAGC 657

Qy 177 ArgGluHisLeuAlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIle 196
Db 658 CGAGAGACCTCAAGTGAAGCAAGTATTGCTGCGACAGACAGCCTCTTGAGAAATA 717

Qy 197 ValGluPheHisAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeu 216
Db 718 CTGAATTCATCAGAGACAGTGGCCAGACACTGCTGAGTCCGATTTCCAGTGTCTC 777

Qy 217 GluIleAlaArgAlaGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGlu 236
Db 778 GAATTTGCTGAAAGTTGAAATGTCAGCATCGATTTTCACTAGCTTTTGACAGGAAA 837

Qy 237 GlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
Db 838 GAAACCAAGATTCACTGCGAGTTTCCACATGGGTGTACTCGTGTCCAGGGACACAC 897

Qy 257 LysIleAsnAlaPheAsnThrAlaLysValArgLysLeuSerPheLysValGlyPhe 276
Db 898 AAATATCAACTTTCATCGTCCAGAGTCCGTAACCTTAAGCTTCAAGAGAAAGATTT 957

Qy 277 LeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
Db 958 CTTATCAAACTTTCATCCAGAGGTTTCATGACCTTACAGACACATTTAGATTTTGTG 1017

Qy 297 AlaSerArgAspPheCysLysSerPheTyrLysIleCysValGluHisAlaPhePhe 316
Db 1018 GGATGATGAGATGAATGTAAGAACTTCTGGAAGATTTGTGTGAGATTCACACTTTT 1077

Qy 317 ArgLeuPheGluGluProLysProLysProLysProValLeuPheSerArgLysSer 336
Db 1078 AGACTTTTGACCAACTTAAGCCAAAGCCAAAGCCGCTTCTTCAAGCCGGGCTCTCTCC 1137

Qy 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyHis 356
Db 1138 TTGAGATACAGTGAAGAACTCGAAACAACTGTGATTTATTTCAAGACAGTGAATG 1197

Qy 357 LysLeuValGlnPheGluArgLysHisSerLysIleHis---SerIleArgSerLeu-- 374
Db 1198 AAGAGATTCATATGAAAGAGACAGAGAACCCACAGTCCGTTGAGCTTGACT 1257

Qy 375 AlaSerGlnProThrLeuAsnSerGluValLeuGluGlnSerGlnGlnSerThrSer 394
Db 1258 GCAGACCTACCAAAAG-----AGC 1278

Qy 395 LeuThrPheGlyGluIleArgLysSerProGlyGlyGlnSerCysArgArg----- 411
Db 1279 ATCTCATTTCCCGAGGAGATTGAGAGACTCTCTCCCTCCCATCTTCAGAGAAATGCTTTAC 1338

Qy 412 ---GlyLysGluProLysValSerAlaGlyLysProGly----- 423
Db 1339 TCGCTCTCCCTCCACTCTGCTCCCTGCTGCGCTCCAGAGTTTAAAGACAGACAGC 1398

Qy 424 ---SerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAlaAsp 441
Db 1399 TCCCTCACAGATCCCGAGTTTCTTACGTCAAGAGTCCAGCTGACAGAGCGCAGTGA 1458

Qy 442 GlyAlaAlaSerAlaProThrGlnGluGluGluValValLysAspArgThrGln 461
Db 1459 GCAGTGGCTGAGGCCCCAG-----AACCATCG 1488

Qy 462 SerLysProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
Db 1489 GCCCAGCCCTCGGGCCCCCGCATCCAGCTGTGTCAGGCTTTCCAGAAAGTCC 1548

Qy 479 HisLeuSerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeu 498
Db 1549 CAGCTTCTCCTCCAGCCGGAAGACCCCTGAGTGTGAGCTTCA-----TTT 1599

Qy 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuLeuLeuSerProLeuAsn 518
Db 1600 CAGTCCCTTTGGGCCCACTGAAACAGGGCTCATCCCACTCTGAGCCCTGTCTCAGT 1659

Qy 519 AspGlnAlaCysProArgThrAspAspGluAspGluGlyArgArgLysPheProThr 538
Db 1660 GATGCTGGCGGAGCCGGATGACTCGAG---GAGCCAGACACAAAGCCGCTGCA 1716

Qy 539 AspLysAlaTyrPheIleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAsp 558
Db 1717 GACGAGGCTTACTTCAATAGTCAAGAGATTCTCGCTACAGAAACATACCTCAAGAT 1776

Qy 559 LeuGluValIleThrSerTyrPheGlnSerThrValSerLysGluAspAlaMetProGlu 578
Db 1777 TTGAGATTTATTCAGTGTGTTCCGACAGCGAGTGTGAAGAGAGACCGCATGCCCTGCG 1836

Qy 579 AlaLeuLysSerLeuIlePheProAsnPheGluProLeuHisValPheHisThrAsnPhe 598
Db 1837 ACTCTATAGAGCTGCTCTTCTTCCAACTGATCCCATGATAGTTTCCACAGAGCTTC 1896

Qy 599 LeuLysGluIleGlnGlnArgLeuAlaLeuTyrGlnGlyArgSerAsnAlaGlnIleArg 618
Db 1897 CTGCGGAGGTGAGAGAGGCTGGCACTGTGAAAGGGCCCTCCAAAGCCACACAAAA 1956

Qy 619 ---AspTyrGlnArgIleGlyAspValMetLeuLysAsnIleGlnLysCysHisLeu 637
Db 1957 GCGAGTCATCAAGATCCGGGACATCCGCTCAGAAACATGGCCAGTTAAAGAGTTT 2016

Qy 638 AlaAlaHisLeuThrLysHisSerGluAlaLeuGluAlaLeuGluAsnGlyIleLysSer 657
Db 2017 ACCAGCTACTTCCAAAGACATGACGAGTCTTAAAGAACTGGAAGAAAGCTACCAAGC 2076

Qy 658 SerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuPro 677
Db 2077 TGTAAAGATGTGAGGCACTGTACAAAGAGTTTGAAGCTGCAAGAGGTCTGCTTGCCT 2136

```

Qy 678 LeuAnThrPheLeuLeuArgProLeuH1sArgLeuMeth1sTyrTyrLeuValLeuGlu 697
Db 2137 CTCACACGCTTCCTGCTAAACCCATCCAGCGGCTGCTGCACTACCGCTGCTGGCC 2136
Qy 698 ArgLeuCyVlyVh1sVh1sProProSerH1sAlaPheArgArgCyVArgAlaAlaLeu 717
Db 2197 CGCCTATCGGACATTAACAGCCCGGACCATGACTACGCTGACTGCGCATGACGCCCTG 2256
Qy 718 AlaGluIleThrGluMetValAlaGluLeuH1sGlyThrMetC11elyMetGluAsnPhe 737
Db 2257 AAGCCATCACAGAGTGACCAACCACTACAGCACTATTCATCCGCTGAGAACTCG 2316
Qy 738 GlnIlyLeuH1sGluLeuLyLeuAAsPheU1leGly11eAAsPheValValProGly 757
Db 2317 CAGAACTAAAGGAGCTGACAGCGGACCTGGTGGCATAGAACTCATGCTCTGGC 2376
Qy 758 ArgGluPhe11eArgLeuGlySerLeuSerLyLeuSerGlyLyGlyLeuGlnGlnArg 777
Db 2377 AGGAGATTCACTCGTAGGGCTGCTTCAAGAGCTCAACAAAGAGGCTGACAGAGG 2436
Qy 778 MetPheLeuPheAsnArgValLeuLeuTyrThrSerArgLyLeuThrAlaSerAsn 797
Db 2437 ATGTTTTCCTGTTCTAGATATGTTGCTGTACACAAAGCAAGAGTTGACAGGAGCCAGC 2496
Qy 798 GlnPheLyValH1sGlyGlnLeuProLeuTyrGlyMetThr11eGluGluSerGluAsp 817
Db 2497 CACTCCGGATCCGGGCTCTTCCCTCCAGGCACTGCTGGTGAAGAAAGATATAC 2556
Qy 818 GluTyrGlyValProH1sCyVleuThrLeuArgGlyGlnArgGlnSer11e1ValAla 837
Db 2557 GAGTGCTCTGTTCCACACTGTTTCAACCATTCACGCGCTCAAAACAAATGCTGTGGCA 2616
Qy 838 AlaSerSerArgSerGlyMetGluTyrValGluAsp11eGlnMetAla11eAspLeu 857
Db 2617 GCCAGCACTCGCTGGAAGAAAGAGAGTGTGACCTCGCAATCCGCGCATCCAAAGCA 2676
Qy 858 AlaGluLySerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
Db 2677 GCCAAGAGTGGCGGTGACAGGCCCCCTGCACTGCCAGGCGGCACGTGTGCACTGCTGCC 2736
Qy 874 ProPheAsnLySerProAspArgLyAlaThrAlaAlaAspGlnIlySerGluAspAspLeu 893
Db 2737 CCC-----AGATCCCCCAAGAG---GTATCTGTGAGAGAGTCAAGAAATGATGCT 2787
Qy 894 SerAlaSerArgThrSerLeuGluArgGlnAlaPheH1sArgLyLeuThrMetValH1s 913
Db 2788 CGGGGTGTCGCGACCTCCCTGGAAGGAGCATGCGCAACCGGGCAACCAACATGAC 2847
Qy 914 ValCyATrPnH1sArgAsnThrSerValSerMetValAspPheSer11eAlaValGluAsn 933
Db 2848 GTGTGCTGTGTAACCGAAACAGAGCTGTCCAGGGCAACACACATGACGTGTGAGAAC 2907
Qy 934 GlnLeuSerGlyLeuLeuLyArgLyPheLyAsnSerAsnGlyTyrGlnLyLeuTyr 953
Db 2908 CAGCTTTCAGGATATCTGTAAAGAAAGTTCAAAACAGTATGCTGGCAGAAAGCTCTGG 2967
Qy 954 ValValPheThrAsnPheCyVleuPhePheTyrLySerH1sGlnAspAsnH1sProLeu 973
Db 2968 GTGCTTTTAACTCAATCTGTGTGTTCTTCAAAATCATCAAGATGATACCACTG 3027
Qy 974 AlaSerLeuProLeuLeuGlyTyrSerLeuThr11eProSerGluSerGluAsn11eGln 993
Db 3028 GCCAGCTTCCCGCTGCGGTGACAGCTGACATCCCAAGGAGCGCGATGCAATAC 3087
Qy 994 LyAspTyrValPheLyLeuH1sPheLySerH1sValTyrTyrPheArgAlaGluSer 1013
Db 3088 AAGAGCATATGTTTCAAGCTCAGTCAAAATCCACAGCTCACTTCCGGGTGAGAGC 3147
Qy 1014 GluTyrThrPheGluArgTyrMetGluVal11eArgSerAlaThrSerSerAlaSerArg 1033
Db 3148 AAGTACATTTTGAAGGTGAGTGAAGTATCCAGGGGCGCAACAGCTCAGCGGAGG 3207

```

```

ADRs5675
ID ADRs5675 standard; DNA; 3997 BP.
XX
XX ADR25675;
XX
XX 21-OCT-2004 (first entry)
XX
XX Breast cancer prognosis marker #1536.
XX
XX ds; breast cancer; prognosis; gene expression; diagnosis.
XX
XX Homo sapiens.
XX
XX MO2004065545-A2.
XX
XX 05-AUG-2004.
XX
XX 15-JAN-2004; 2004WO-US001100.
XX
XX 15-JAN-2003; 2003US-00342887.
XX
XX (ROSE-) ROSETTA INPHARMATICS LLC.
XX
XX (MECA-) NETHERLANDS CANCER INST.
XX
XX Van't Veer LJ, He Y;
XX
XX WPI: 2004-593473/57.
XX
XX Classifying a breast cancer patient according to prognosis comprises
XX determining the similarity between the level of expression of each of
XX five genes in a cell sample taken from patient, to control levels.
XX
XX Disclosure; SEQ ID NO 1536; 226bp; English.
XX
XX The invention relates to a method of classifying a breast cancer patient
XX according to prognosis by determining the similarity between the level of
XX expression of each of five genes for which markers are listed in the
XX specification, in a cell sample taken from the breast cancer patient, to
XX control levels of expression for each respective five genes to obtain a
XX patient similarity value. The methods are useful for classifying a breast
XX cancer patient according to prognosis. Kits and computer program products
XX are useful for data analysis using the diagnostic, prognostic and
XX statistical methods of the invention. This sequence corresponds to a
XX marker used in the method of the invention.
XX
XX Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1,31e-189 Length: 3997
XX Score: 2941.50 Matches: 580
XX Percent Similarity: 70.0% Conservative: 162
XX Best Local Similarity: 54.7% Mismatch: 261
XX Query Match: 13.8% Indels: 57
XX DB: 13 Gaps: 14
XX
XX US-09-555-342B-2 (1-1045) x ADR25675 (1-3997)
Qy 1 MetGlyGluIleGluGlnArg-----ProThrProGlySerArgLeuGlyAla 16
Db 118 ATGGGGGAGATAGAGAAAGAAATACAGAGTCTGCAAGCTCAGGGAGTGGCTGGGCGCC 177
Qy 17 ProGluAsnSerGly11eSerThrLeuGluArgGlyGlnLyAspProPheProThrProSer 36
Db 178 CAGAACCCCTGTGGAGTTAGCAACCTTGAACCTGGGCAAGCTCTTGGCCAGAAATGCAA 237
Qy 37 GLyLySerValSer11eLyVal11eGlnMetLeuAspAspThrGlnGlnAlaPheGluVal 56
Db 238 GAGAAAGCACTGCAACCTCAAGATTAAGCTCTGGAACCAACCAATGAAATATTGACATT 297
Qy 57 ProGlnArgAlaProGlyLyValLeuLeuAspAlaValCyAsnH1sLeuAsnLeuVal 76
Db 298 GAGCTTAATGCGATGGCCAGATGATTAATGACACAAGTGTGAAAGCGTTTAAACCTGGTA 357

```

QY	77	GIUGIYAPRYrPhneGlyLeuGluPheProAspHisblyblysllethnValtrPLeuAsp	96
Db	358	GAATGTGACTCTCTCGGATGGAGTTTAAAATATCTCGTCTACGTGATTTGGCTTGA	417
QY	97	LeuLeuLysProIleValIyGlnIleArgArProLysHisValValIValIyPheVal	116
Db	418	CTTATGAAACCATATATTAGGCMAATACAAAGGCCAAAGAAATGTGTGCTTCGGCTAGCT	477
QY	117	ValIyPhePheProProAspHisbThrgInIeugInIuGluLeuThrArgTyrLeuPhe	136
Db	478	GTAAATTTTTTCCACTGTATCTGTGTGACTACAAAGAAAGATATACAGATATCTTGT	537
QY	137	AlAlEugInValIyGlnAspLeuAlaGlnIlyArgLeuThrCyAsnAspThrSerAla	156
Db	538	GCCTTGCACCAATTACAGAGACCTGTGGAAAGACCTTGTGACCTGTGTGACACCAACGC	597
QY	157	AlAlEuleuLleSerHisIleValGInSerGluIleGlyAspPheAspGluAlaLeuAsp	176
Db	558	GGCTTCTCAGTCCCATCTCTGTGAGTCCGGAAATGAGATATTCAGATGAACCTGGAC	657
QY	177	ArgGluHisLeuAlaIyAsnLysTyrIleProGInIAspAlaLeuGluAspLysIle	196
Db	658	CGAGAGCACTCAAGTAGACGAGATTTGGCTGGCCAGAGCACTGCTTGAAGAAGTA	717
QY	197	ValGluPheHisAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuLeu	216
Db	718	CTAGAAATTCACATCAGAACACCTGTGGCCAGACACTGTGTGATGCGATTCACAGTCTC	777
QY	217	GluIleAlaArgArGluGluMetTyrGlyIleArgGluHisProAlaIyAspArgGlu	236
Db	778	GAATATGCTCGAAAGTGTGAATGTATCGGATTCAGATTTCAATGAGCTTCTGACAGGGAA	837
QY	237	GlyThrLysIleAsnLeuAlaValAlaAsnThrgIyIleLeuValPheGlnIyPheThr	256
Db	838	GGAAACAAGATTCACTGGCAGTTTCCACATGGGTATCTGTGTTCAGAGGACCAACC	897
QY	257	LysIleAsnAlaPheAsnTrpAlaLysValArgIyLysSerPheLysArgLysPhe	276
Db	898	AAATCAACACTTTCACATGGTCCAGAGCTCGTAACTTCAACTTCAAGAGGAAAGATTT	957
QY	277	LeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspTrpLeuGluPheLeuMet	296
Db	958	CTTATCAAACTTCATCCAGAGGTTATGACCTTACAGACACATTAAGATTTTGTG	1017
QY	297	AlAspSerArgAspPheCybLysSerPheTrpLysIleCybValGluHisAlaPhePhe	316
Db	1018	GGTATTAAGATGATGTATGAAGACTTCGGAAGATTGTGTGAGATACACACTTTTTT	1077
QY	317	ArgLeuPheGluGluProLysProLysProLysProValLeuPheSerArgLysSer	336
Db	1078	AGACTTTTGGACCAACTTAAGCCAAACCAAAAGCCGTCCTTCCAGCCGGGCTCTCC	1137
QY	337	PheArgPheSerGlyArgThrGlnIyGlnAlaLeuAspTyrValIyGlnGlyHis	356
Db	1138	TTCAATATCACAGTGAAGAACTCAGAAACACTAGTATGATTTTCAAGACAGTGAATG	1197
QY	357	LysLysValGlnPheGluArgLysHisSerLysIleHisSerIleArgSerLeu	374
Db	1198	AAGAAATATTCATATGAAAGAAAGCAKAGCAAGAACCCACAGTCTCGTCCAGCTGACT	1257
QY	375	AlAspSerGlnProThrGluLeuAsnSerGluValIeugInIuSerGlnIuSerThrSer	394
Db	1258	GCAGACCTACCAAAACAG-----AGC	1278
QY	395	LeuThrPheGlyGlnGlyAlaGluSerProGlyGlyGlnSerCybArgArg-----	411
Db	1279	ATCTCATTTCCCGAGAGGATGAGACTCTGCTCCCATCTTACAGCAATGCTTTTAC	1338
QY	412	-----GlyLysGluProLysValSerAlaGlyIuProGly-----	423
Db	1339	TGCGTCTCTCCCTCCACTCTGTGTCCCTCTGGCTCCAGAGATTTAAGACAGACAGAC	1398
QY	424	-----SerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAlaAsp	441

[illegible]

Db 2437 ATGTTTTTCTGTCTCATGATATGTTGCTGTACACAAAGAGATTGACGAGCACC 2496
 Qy 798 GlnhlyeValhAglyGlnleuProleuTyrglyMetThrllleGlnGlnSerGluAsp 817
 Db 2497 CACTTCGGGATCCGGGGCTCTCTCCCTCCCAAGGACGTGGTGAAGAAAGTATAC 2556
 Qy 818 GlnTyrglyValProhAsCyLeuThrhleuArgglyGlnArgGlnSerllleValAla 837
 Db 2557 GAGTGGTCTGTTCCACACTGTTTCCACATTCACGGGGCTCCGAAAACATGTGTGCA 2616
 Qy 838 AlaSerSerArgSerGluMetGlnlyeTyrglyValGlnAspIleGlnMetAlaIleAspLeu 857
 Db 2617 GCCAGCACTCGGCTGGAGAAAGAAAGTGAAGTGGACCTGAACCTCCGGCATCCAAACA 2676
 Qy 858 AlaGlnlyeSerSerSerProAlaProGluPhe-----leuAlaSerSerPro 873
 Db 2677 GCCAAGAGTGGCGGTGACACGACGCTGACCTGACCTGCCAGGCCCACTGTGTGACCTGCC 2736
 Qy 874 ProhAsnlyeSerProAspGlnAlaThrAlaIleAspGlnGlnSerGluAspAspLeu 893
 Db 2737 CCC-----AGATCCCCCAAGAG--GTATCTCGAGCGAGAGTCCAGAAAGATGATCT 2787
 Qy 894 SerAlaSerArgThrSerLeuGlnArgGlnAlaProhIleArgGlyAsnThrMetValhIs 913
 Db 2788 CGGGGTGTCGACAGTCCCTGAGAGGGGATGGCCAGCACCGGGCCAAACACCAATGCAC 2847
 Qy 914 ValCystrpHisArgAsnThrSerValSerMetValaAspSerIleAlaValGlnAsn 933
 Db 2848 GTGTGCTGTGACCGGAACACACGAGGTGTCCAGGGCAGACCCAGTGCAGCTGTGAGAAC 2907
 Qy 934 GlnleuSerGlyAsnleuLeuAlaArglyPhehlyeAsnSerAsnGlyTyrglyGlnleuTy 953
 Db 2908 CAGCTTTCAGGATATCTGCTTAAGAAAGTTCAAAACAGATCATGGCTGGCAAGAGCTTGG 2967
 Qy 954 ValValPheThrAsnPheCyLeuPhePheTyrlLySerSerhIsGlnAspAsnIleProleu 973
 Db 2968 GTGCTCTTACCAACTCTGTTGTTCTTCAAAACTCATCGAGTCACTACCACTG 3027
 Qy 974 AlaSerleuProleuLeuGlyTyrlSerleuThrIleProSerGlnSerGluAsnIleGln 993
 Db 3028 GCCAGCTCCCGCTGGCTGGGTACAGAGCTGAGCATCCCCAGGAGGCCGATGGCATACAC 3087
 Qy 994 LysAspTyrglyValPhehlyeLeuHIsPhehlyeSerhIsValTyrglyPheArgAlaGlnSer 1013
 Db 3088 AAAGACTATGTTTTCAGAGCTCAAGTTCAAATCCAGCTTCTTCCGGGCTGAGAGC 3147
 Qy 1014 GlnTyThrPheGlnArgTyrglyMetGlnValIleArgSerAlaThrSerSerAlaSerArg 1033
 Db 3148 AAGTACACATTTGAAGGTGATGAGGTGATCCAGGGGGCCAGCAGCTCAGCCGGAGG 3207
 RESULT 8
 ADS64452 standard; DNA; 3997 BP.
 AC ADS64452;
 XX 16-DEC-2004 (first entry)
 DT
 XX Human KIAA0793 gene.
 DE
 XX Human; Interferon; IFN; gene; ds; KIAA0793.
 KM
 XX Homo sapiens.
 OS
 XX US2004185489-A1.
 PN
 XX 23-SEP-2004.
 PD
 XX 17-MAR-2004; 2004US-00802432.
 PF
 XX 21-MAR-2003; 2003BP-00006263.
 PR
 XX (CERT/) CERTA U.

PA (FOSE/) FOSER S.
 PA (WEYE/) WEYER K.
 XX Certa U, Foser S, Weyer K;
 PI
 XX MPI; 2004-689187/67.
 DR GENBANK; AB018336.
 XX
 PT Determination of biological activity of compound which can modulate gene
 PT transcription, involves contacting host with compound, determining
 PT general transcriptional gene response of host, and quantitating gene
 PT response induced by compound.
 XX
 PS Claim 8; SEQ ID NO 16; 53pp; English.
 CC
 CC The invention relates to a method for determining the biological activity
 CC of a compound which can modulate gene transcription. The method involves
 CC contacting a host with a compound, determining the general
 CC transcriptional gene response of the host and quantitating the gene
 CC response induced by the compound. This gene transcription assay method is
 CC useful for the determination of the biological activity of a compound
 CC which can modulate gene transcription. The present sequence is the human
 CC KIAA0793 gene. This sequence is used to estimate the influence of the
 CC pegylation site on the transcriptional activity of interferon (IFN) by
 CC measuring this IFN-induced gene expression pattern of the melanoma cell
 CC line ME15 via the oligonucleotide array technology.
 XX
 SQ Sequence 3997 BP, 998 A, 1132 C, 1068 G, 799 T; 0 U, 0 Other;
 Alignment Scores:
 Pred. No.: 1,31e-189 Length: 3997
 Score: 2941.50 Matches: 580
 Percent Similarity: 70.0% Conservative: 162
 Best Local Similarity: 54.7% Mismatches: 261
 Query Match: 53.8% Indels: 57
 DB: 13 Gaps: 14
 US-09-555-342B-2 (1-1045) x ADS64452 (1-3997)
 Qy 1 MetGlyGlnIleGlnGlnArg-----ProThrProGlySerArgLeuGlyAla 16
 Db 118 ATGGGGAGATAGAGAGAACATACAGAGTCTCGACAGCTGACAGGAGATGGCTGTGGTGC 177
 Qy 17 ProGluAsnSerGlyIleSerThrLeuGlnArgGlyGlnlyeProProProThrProSer 36
 Db 178 CAGACCCCTGTGGAGTTAGACCCCTGAGCTGGGAGACTCTTTCGCCAGAAATGCA 237
 Qy 37 GlnlyeLeuValSerIleLeuIleGlnMetLeuAspAspThrGlnGlnAlaPheGlnVal 56
 Db 238 GAGAAAGCACTGCACCTCCAGAGTAAAGCTGTCGACAAACCAAGAAATTTTGACATT 297
 Qy 57 ProGlnArgAlaProGlyLyValleuLeuAspAlaValCyAsnHIsleuAsnleuVal 76
 Db 298 GAGCTTAAGTCCGATGCCAGGTATTTACTGACCAAGATGGAGAGCCGTTTAAACCTGGTA 357
 Qy 77 GlnGlyAspTyrglyPheGlyLeuGlnPheProAspHislyeLyIleThrValIleTyrLeuAsp 96
 Db 358 GAATGTACTACTTCCGAGATGAGAGTTTCAAAATCTAGCTTCACTGATTTGGCTTGAA 417
 Qy 97 LeuTyAspProIleValIyGlnIleArgArgProhlyeHisValIleValIyPheVal 116
 Db 418 CCTATGAAGAACCATTCATTAGCAAAATACGAAGGCCAAAGATGTGTGCTTCGCTAGCT 477
 Qy 117 VallyePhePheProProAspHisIsthGlnleuGlnGlnGlnleuTyThrArgTyrlLeuPhe 136
 Db 478 GTAAATTTTTCACCTGATCTCTGTACAGTACAAAGAAATATACAAAGATCTTGT 537
 Qy 137 AlaGlnGlnVallyeGlnAsnleuAlaGlnGlyArgleuThrCyAsnAspThrSerAla 156
 Db 538 GCTTGCACTTAAAGAGACCTGCTGGAAAGGCTTTGACCTGTGTGACACACAGCG 597
 Qy 157 AlaLeuLeuIleSerhIsIleValGlnSerGlnIleGlyAspPheAspGlnAlaLeuAsp 176

Db 598 GCCCTTCTCAGCTCCCATCTCTGCACTGCGAAATAGAGATTACGATGAACGCTGAC 657
 Qy 177 ArgGlnHisLeuAlaValSerLysTyrIleProGlnGlnSerAlaLeuGluSerPheProthr 196
 Db 658 CGAGAGACCTCAAGTAGAGATATTGCTGGCCAGACGACCTGCTTGAAGATA 717
 Qy 197 ValGluPheHisValSerHisIleGlyGlnThrProAlaGluSerAspPheGlnLeu 216
 Db 718 CTGAATTCATCAGAGACGCTGGCCAGACCTGCTGAGTTCGATTCCAGGTCTC 777
 Qy 217 GluIleAlaArgGluMetTyrGlyIleArgLeuHisProAlaValSerAspGlu 236
 Db 778 GAATGCTCGAATGGAAATGTACGGCATCAGATTCCATGCTCTGACAGGAA 837
 Qy 237 GlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
 Db 838 GGAACCAAGATTCACTGGCAGTTTCCACATGGGTACTCGTGTTCAGGGCACACC 897
 Qy 257 LysIleAsnAlaPheAsnThrAlaValAlaArgLysLeuSerPheLysArgLysArgPhe 276
 Db 898 AAATCAACACTTCAACTGCTCCAGTCCGTAATTAAGCTTCAAGAGAAAGATT 957
 Qy 277 LeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
 Db 958 CTATCAACTTCATCCAGAGGTTCACTGACCTTACCCAGACACATTGAATTTGTTG 1017
 Qy 297 AlaSerArgAspPheCysLysSerPheThrLysIleCysValGlnHisValAlaPhePhe 316
 Db 1018 GGTAAGTAGAGTAGATGTAAAGACTTGAGAGATTGTGTGGATATCAACCTTTT 1077
 Qy 317 ArgLeuPheGluGluProLysProLysProLysProValLeuPheSerArgLysSer 336
 Db 1078 AGACTTTGGACCACTTAAGCCAAAGAAAGCCGCTTCTTCAAGCGGGCTCTCC 1137
 Qy 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyHis 356
 Db 1138 TTCAAGTACAGTGAAGAACTCAGAAACAATGATATTATTCAAGACGTGATG 1197
 Qy 357 LysLeuValGlnPheGluArgLysHisSerLysIleHis---SerIleAspSerLeu--- 374
 Db 1198 AAGGAATTCATATTAAGAAAGGACAGACAGAACCCACAGCTCCGTTCAAGCTTGACT 1257
 Qy 375 AlaSerGlnProThrGluLeuAsnSerGluValLeuGluGlnSerGlnSer 394
 Db 1258 GCAGACTTACCAAAACAG-----AGC 1278
 Qy 395 LeuThrPheGlyGluAlaGluSerProGlyGlyGlnSerCysArgArg----- 411
 Db 1279 ATCTCATTTCCCGAGGATGAGAGACTGCTCCCATCTTCAAGGAATGCTTTTAC 1338
 Qy 412 ---GlyLysGluProLysValSerAlaGlyGluProGly----- 423
 Db 1339 TCGCTCTCTCCCTCCACTGCTGCTCCCTGCTGCTGCAAGTTTAAGACAGACAGAC 1398
 Qy 424 ---SerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAlaAsp 441
 Db 1399 TCCCTCAGACATCCCGAGGTTTCTCTACGTCAGAGCTCAGACAGAGAGCCAGTGA 1458
 Qy 442 GlyAlaAlaSerAlaProThrGluGluGluGluValValLysAspArgThrGln 461
 Db 1459 GCAGTGGCTGAGGCCCGAC-----ACACCATCG 1488
 Qy 462 SerLysProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
 Db 1489 GCCCAGCCCTCGGGGCCCGCCGACCTCAAGCTGCTGCAAGCCCTTCCAGAAAGTCT 1548
 Qy 479 HisLeuSerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeu 498
 Db 1549 CAGCTTCTCCCTCCAGCCGGAAGAGCCCTGAGTGTGAGCCCTGCA-----TTT 1599
 Qy 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuLysSerProLeuAsn 518
 Db 1600 CAGGTGCTTTGGGCCAGCTGAACAGGCTCATCTCCCACTCGAAGCCCTGTCTCAGT 1659

Qy 519 AspGlnAlaCysProArgThrAspArgLysAspGluGlyValArgArgLysArgPheProthr 538
 Db 1660 GATGCTGGGGAGACCCGGAGTGACTGCGAG---GAGCCAGACACAAAGCCGCTCGCA 1716
 Qy 539 AspLysAlaTyrPheIleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAsp 558
 Db 1717 GACGAGCCCTACTTCAATGACAAAGAAATTTCTGCTACAGAAAGAACTATCTCCAAAGAT 1776
 Qy 559 LeuGluValIlePheSerThrPheGlnSerThrValSerLysGluAspAlaMetProGlu 578
 Db 1777 TTAGAATTATTAACCTGTGTTCGCGACGCGAGTGTGAAGAGAGACGCTGCGG 1836
 Qy 579 AlaLeuSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAsnPhe 598
 Db 1837 ACTGTATGACGCTGCTCTTCCAAATCGATCCCATGTATGATTCACAGAGCTTC 1896
 Qy 599 LeuLysGluIleGluGlnArgLeuAlaLeuThrGluGlyArgSerAsnAlaGlnIleArg 618
 Db 1897 CTGCGGAGGTGGAGCAGAGCTGCGCATCTGGGAAGGCCCCCTCCAAAGCCACACAAA 1956
 Qy 619 ---AspTyrGlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysIleu 637
 Db 1957 GGCATCATCAAGATCGAGGACATCTGCTCAGGAACATGCGCAATTAAAGAGTTT 2016
 Qy 638 AlaAlaHisLeuThrLysHisSerGluAlaLeuGluAlaLeuGluAsnGlyIleLysSer 657
 Db 2017 ACCAGCTACTCCAAAGACATGACGAGTCTTACAGAACTGGAAGAAAGCTACCAAAAGC 2076
 Qy 658 SerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuPro 677
 Db 2077 TGTAAGATGTGGAGCAGTGTACAGAGTTTGACTCAGAAAGCTGCTGCTTGGCT 2136
 Qy 678 LeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisArgTyrGlnValLeuGlu 697
 Db 2137 CTCAACAGTTCTGCTGTAAAGCCATCCAGGCGCTGTGACTTACGCGCTGCTGCGC 2196
 Qy 698 ArgLeuCysLysHisValProProSerHisValAspPheArgAspCysArgAlaAlaLeu 717
 Db 2197 CGCTATGGGACATTAACAGCCCGGCGACACATGACTAGCGTGAAGCCATGACGCGCTG 2256
 Qy 718 AlaGlnIleThrGluMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPhe 737
 Db 2257 AAGCCATCAGAGGCTGACCAACACATCAAGCACTTCTCATCCGGTGGAGAACTTG 2316
 Qy 738 GlnLysLeuHisGluLeuLysValAspLeuIleGlyIleAspAsnLeuValProGly 757
 Db 2317 CAAAGCTAAAGAGCTGCGAGCGGACCTGTGTGGCATTAAGAACTCATTGTCTCTGGC 2376
 Qy 758 ArgGluPheIleArgLeuGlySerLeuSerLysValSerGlyLysGluGlnArg 777
 Db 2377 AGGGAATTCATCCGTGAGGGCTGCTTCAAGCTCACCAAGAAAGGCGCTGACAGAGG 2436
 Qy 778 MetPhePheLeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsn 797
 Db 2437 ATGTTTTTCTGTTCACATATATGTGTCGTGCACAAAGAAAGAGTGTGAGGAGCAGC 2496
 Qy 798 GlnPheLysValHisGlyGlnLeuProLeuTyrGlyMetThrIleGluGluSerGluAsp 817
 Db 2497 CACTTCGGAATCGGGGCTCTTCTCCCTCCAAAGGAGTGTGTGTGAAGAAATGATTAAC 2556
 Qy 818 GluThrGlyValProHisCysLeuThrLeuArgGlyLysArgGlnSerIleIleValAla 837
 Db 2557 GAGTGTCTGTTCCACACTGTTTCAACATCTACGCGGCTCAGAAACAAATCGTGTGCA 2616
 Qy 838 AlaSerSerArgSerGluMetGlyLysThrValGluAspIleGlnMetAlaIleAspLeu 857
 Db 2617 GCCAGCATCTCGGTGGAAGAAAGTGAATGTAAGTCTGAACCTCCGGAATCCAGCA 2676
 Qy 858 AlaGluLysSerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
 Db 2677 GCCAAGTGGCGGTGACAGGCGCTTGCATCCAGCGCGCATGTGTGCACTGCTGCC 2736

QY 874 ProkaphenylSerProaspGluAlaThrAlaIaaspGlnIuSerGluaspLeu 893
DB CCC-----AGATCCCCCAAGAG---GTATCTCGAGCAGGAGTCAGAAAGATGCT 2787
QY 894 SerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgGluAsnThrMetValHis 913
DB CGGGGTGTCGACAGCTCCCTGAGAGGGGATGCGCAGACCGGGGCAACACCAATGCAC 2847
QY 914 ValCysThrPheValArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
DB GTGTGCTGGAGACCGGAACACGAGGTGTCCAGGCAAGCCACAGTCAGCTGTGAGAAC 2907
QY 934 GlnLeuSerGluAsnLeuLeuArgLysPheLysAsnSerAsnGlyTyrGlnLysLeuTyr 953
DB CAGCTTCAGAGATATCTGCTTAAAGAACTTCAAAACAGTCATGCTGCGCAGAACCTTGG 2908
QY 954 ValValPheThrAsnPheCysLeuPhePheTyrLysSerHisGlnAspAsnHisProLeu 973
DB GTGCTCTTACCAACTCTGTGTGTCTTCTTACAAAATCATCATCAGATGATACCACTG 3027
QY 974 AlAserLeuProLeuLeuGlyTyrSerLeuThrIleProSerGluSerGluAsnIleGln 993
DB GCCAGCTCCCGCTGGCTGGCTGACGATGCCCGAGGAGCCGATGGCATCAC 3087
QY 994 LysAspTyrValPheLysLeuHisPheLysSerHisValTyrTyrPheArgAlaGluSer 1013
DB AAGACATATGTTTCAAGCTCCAGTCAATCCAGCTTCTTCCGGGCTGAGAGC 3147
QY 1014 GluTyrThrPheGluArgTyrMetGluValIleArgSerAlaThrSerSerAlaSerArg 1033
DB AAGTACATTTGAAAGATGATGAGATGATGATGAGGAGGCGCAGACCTCAGCCGGAGG 3207

RESULT 9

AAC98992
ID AAC98992 standard; cDNA; 3094 BP.

AAC98992;

09-MAR-2001 (first entry)

Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:220.

Human: pancreas; pancreatic cancer; pancreatic cancer antigen; detection;
diagnosis; identification; cytotoxic; neuroprotective; nootropic;
immunomodulatory; relaxant; contraceptive; gynaecological;
antiinflammatory; cardiant; gene therapy; chromosome mapping;
linkage analysis; tissue identification; tissue typing; forensic; neutral;
immune system; muscular; reproductive; gastrointestinal; pulmonary;
cardiovascular; renal; proliferative; ss.

Homo sapiens.

MO200055320-A1.

21-SEP-2000.

08-MAR-2000; 2000MO-US005989.

12-MAR-1999; 99US-0124270P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI; 2000-579444/54.

P-PSDB; AAB54227.

New nucleic acid that is a pancreatic cancer antigen for preventing,
treating, or ameliorating a medical condition, particular pancreatic
cancer, or for use in assays for diagnosing a pathological condition.

Claim 1; Page 664-665; 1379pp; English.

CC AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC protein, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention
XX

SO Sequence 3094 BP; 736 A; 849 C; 755 G; 749 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	1.22e-159	Length:	3094
Score:	2498.00	Matches:	479
Percent Similarity:	99.6%	Conservative:	0
Best Local Similarity:	99.6%	Mismatches:	2
Query Match:	45.7%	Indels:	1
DB:	3	Gaps:	0

US-09-555-342B-2 (1-1045) x AAC98992 (1-3094)

QY 565 TrpPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeuLysSerLeuIle 584
DB TGGTTTCAGAGCAGATGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 122
QY 585 PheProAspPheGluProLeuHisLysPheHisThrAsnPheLeuLysGluIleGln 604
DB TTCGGAATTTTGAACCTTGGACAAATTCATACATTAATTTTCAAGGAATTTGAGCAA 182
QY 605 ArgLeuAlaLeuTyrGluGlyArgSerAsnAlaGlnIleArgAspTyrGlnArgIleGly 624
DB CGACTTGGCTTGGAG 242
QY 625 AspValMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaHisLeuTyrLysHis 644
DB GATGTATGCTGAGAGAAATTCAGGGCATAGAGACCTGGCGGCTACCTGTGAGAGCAC 302
QY 645 SerGluAlaLeuGluAlaLeuGluAsnGlyIleLysSerSerArgArgLeuGluAsnPhe 664
DB AGCGAGGCTTGGAG 362
QY 665 CysArgAspPheGluLeuGlnLysValCysTyrLeuProLeuAsnThrPheLeuLeuArg 684
DB TGCAGAGACTTTGAGCTGAG 422
QY 685 ProLeuHisArgLeuMetHisTyrLysGlnValLeuGluArgLeuCysLysHisSerPro 704
DB CAGCTGACCGGCTTACGACTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
QY 705 ProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIleThrGluMetVal 724
DB CCGAGCCAGCGCGACTTCAAGGAGCTCCGAGCGCTTGGCAGAGATCAAGGAGATGAGTG 542
QY 725 AlaGlnLeuHisArgIleThrMetIleLysMetGluAsnPheGlnLysLeuHisArgLeuLys 744
DB GCACACCTCCACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
QY 745 LysAspLeuIleGlyIleAspLeuValValProGluArgGluPheIleArgLeuGly 764

```
DB 603 AAAGATTGATGGCATTCGACATCTTGTTGCCGGGAAGGAGTTTCATCCGCTCGGC 662
QY 765 SerLeuSerIyValSerGlyIyVsgIyLeuGlnGlaArgMerPhePheLeuPheAAsp 784
DB 663 AGCTTAGCAAGCTCTCGGGGAAGGGGCTCCAGCAGAGCTGTTCTTCCTGTTCAAGAC 722
QY 785 ValLeuLeuTyThrSerArgGlyLeuThrAlaSerAsnGlnPheIyValHisGlyGln 804
DB 723 GTCTCTCTATACACGAGCCGGGGGCTGACGGCTCCATCATGTTAAAGTCCAGGGGCAg 782
QY 805 LeuProLeuTyGlyMetThrIleGluGluSerGluAspGluTrpGlyValProHisGys 824
DB 783 CTCCTCCCTCATGTCATGACATTTGAGGAGCAGACGATGGGGGGTCCCTGCG 842
QY 825 LeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaIleSerSerArgSerGluMet 844
DB 843 CTAGACCTCCGGGGCCAGCGGCGAGTCCATCATGCTGCGCCCGCATTCGCTCGAGATG 902
QY 845 GluIyTrpValGluAspIleGlnMetAlaIleAspLeuAlaGluIySerSerSerPro 864
DB 903 GAGAGTGGGTTGAGGACATCCAGATGGCCATTGACCTGGCGGAGAAAGACAGACGCC 962
QY 865 AlaProGluPheLeuAlaSerSerProProAspAsnIySerProAspGluAlaIleTrpAla 884
DB 963 GCCCTGAGATTCTGGCGCAGCAGCCCTCGACAAACAAAGTCCCTGATGAAAGCACCGCG 1022
QY 885 AlaAspGlnGluSerGluAspAspLeuSerAlaSerArgThrSerLeuGlnArgGlnAla 904
DB 1023 GCTGACCAAGAGTCAAGAGATGACTGAGCGCTCGCC-NCATCGCTGAGGCCCGCAGGCC 1081
QY 905 ProHisArgGlyAsnThrMetValHisValCysTrpHisArgAsnThrSerValIleSerMet 924
DB 1082 CGGCACCGGGGCAACCAATGGTGCACGTGTGCTGGGACCGCAACACACGTCCTCAGT 1141
QY 925 ValAspPheSerIleAlaValGluAsnGlnLeuSerGlyAsnLeuLeuArgIyPheIyS 944
DB 1142 GTGAGCTTCGACATCCGACAGTGAAGAAATCACTGTTCTGGAACCTGCTGAGGAATTCGAA 1201
QY 945 AsnSerAsnGlyTrpGlnIyValSerIyValPheThrAsnPheCysLeuPhePheTyR 964
DB 1202 AACAGCAACGGGGGCGAAGAGCTGTGGGTGTTCACAAACTTGTGCTGTTCTTCTAC 1261
QY 965 IySerSerIleGlnAspAsnHisProLeuAlaSerLeuProLeuGlnGlyTySerLeuThr 984
DB 1262 AATTCACACAGGACATCATCTCCCTGGCAGCTGCTGCTCGGCTACTGCTCAGC 1321
QY 985 IleProSerGluSerGluAsnIleGlnIyAspTyValPheIyValHisIyPheIySer 1004
DB 1322 ATCCCTCTGAGTCCGAGAACATCCAGAAAGACTACGTGTTCAAGCTGCACTTCAAGTCC 1381
QY 1005 HisValTyTrpPheArgAlaGluSerGluTyThrPheGluArgTrpMetGluValIle 1024
DB 1382 CAGCTCTACTACTTCAGGGCGGGAAGGAGTACAGTTCAGAAAGGTGAGTGAAGTGAATC 1441
QY 1025 ArgSerAlaThrSerSerAlaSerArgProHisValIleuSerHisIyVsgIuSerLeuVal 1044
DB 1442 CGCAGTGCACCAAGCTCTGCTGCGACCCCAAGTGTGAGCCACAAAGAGTCTCTTG 1501
QY 1045 Tyr 1045
DB 1502 TAT 1504

RESULT 10
ADQ63863
ID ADQ63863 standard; cDNA; 3200 BP.
XX
AC ADQ63863;
XX
DT 07-OCT-2004 (first entry)
XX
DE Novel human cDNA sequence #1024.
XX
KW sb; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
```

```
KW cyostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX
OS Homo sapiens.
XX
PN BP1440981-A2.
XX
PD 28-JUL-2004.
XX
PF 21-JAN-2004; 2004EP-00001196.
XX
PR 21-JAN-2003; 2003JP-00102206.
XX
PR 09-MAY-2003; 2003JP-00131392.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
PI WPI; 2004-535376/52.
DR P-PSDB; ADQ6051.
XX
PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 1024; 2449pp; English.
XX
CC The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.
XX
SQ Sequence 3200 BP; 951 A; 659 C; 718 G; 872 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4.01e-68 Length: 3200
Score: 1150.50 Matches: 265
Percent Similarity: 54.5% Conservative: 89
Best Local Similarity: 40.8% Mismatches: 185
Query Match: 21.1% Indels: 111
DB: 12 Gaps: 15

US-09-555-342B-2 (1-1045) x ADQ63863 (1-3200)
QY 33 ProThrProSerGlyIyLeuValSerIleIyValIleGlnMetLeuAspThrGlnGln 52
DB 161 CCAGAGCGCTCCCTGAGATGATGATCACTTAAAGGCAATTTTGGATGATCCGAGAG 220
QY 53 AlaPheGluValProGlnArgAlaProGlyIyValLeuLeuAspAlaValCysAsnHis 72
DB 221 ATTTTGTGGTGTGATCAAAAGATCATCCGGGAAGGCAATGTTTAACTGAGTGCACCCAT 280
QY 73 LeuAsnLeuValGluGlyAspTyThrPheGlyLeuGluPheProAspHisIySylIeThr 92
DB 281 CTAAATCTTGCTGAAAGAAAGAAATTTTGGATTGAATTTGCAAGCCATTCGAAATTAAT 340
QY 93 ValTrpLeuAspLeuLeuIySerProIleValIyGlnIleArgArgProIyHisValVal 112
DB 341 GTTTGGCTGAGACTTTTGAAGCCATTAACAAAGCAGGTAAATAATCTTAAGAGATTGTT 400
QY 113 ValIyPheValValIyPhePheProProAspHisThrGlnLeuGlnGlnGluLeuThr 132
DB 401 TTCAATTTATGATGAATTTTCCAGTGCAGCCCTGACATCTGCGGAGAAAGAACTTACA 460
QY 133 ArgTyTrpLeuPheAlaLeuGlnValIyVsgIuAspLeuAlaGlnGlyArgLeuThrCysAsn 152
DB 461 AGGTATCTTTTACTCTCAATTAAGAAAGATTGGCTCTGGAAGGCTTCATGCAAGT 520
```

```

QY 153 AepThrSerAlaAlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAsp 172
DB 521 GACAACTGTACAGGCTGATGATCTGCATCTTACATCAAGAACTTGAGACTTTTCAT 580
QY 173 GluAlaLeuAspArgGluHisIleValAspAspValIleProGlnGlnAspAlaLeu 192
DB 581 GAAGAAACAGATAGAGAGCATCTGGCAAACTGGGCTTACCAAAACCAAGACGTGTTA 640
QY 193 GluAspValIleValGluPheHisIleValGlnGlnIleThrProAlaGluSerAsp 212
DB 641 GAGGCAAGATCAAGCACTTTCATCAAGAACATTTGGCAGAGCCGCGAATCTGAC 700
QY 213 PheGlnLeuGlnIleAlaArgArgLeuGlnMetTyrGlyIleArgLeuHisProAla 232
DB 701 ATTGTCTACTGACATAGCAAGAAAGCTGGATATGATGGCATCAGGCTCACCCGCC 760
QY 233 LysAspArgGluGlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPhe 252
DB 761 AGTATGGTGAAGGATGACAGATTCACCTGGCTGTGTCTCATGGAGATCTGTGTGTA 820
QY 253 GlnGlyPheThrLysIleAsnAlaPheAsnThrAlaLysValAlaGlyLeuSerPheLys 272
DB 821 CGGGGAATACAAAGATCAATATCTTTAACTGGGCTAAATCCGCAAGTTGAGTTTAAAG 880
QY 273 ArgLysArgPheLeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeu 292
DB 881 AGAAAGCAATTTCTCATAACTTCATGCCAATATCTTGGTGTGTCAGAAAGATACCTTG 940
QY 293 GluPheLeuMetAlaSerArgAspPheCysLysSerPheThrLysIleCysValGlnHis 312
DB 941 GAGTTCACCATGGCCAGCCAGATGCTGCAAGGCTTCTGGAAGACTGTGTGGAATAC 1000
QY 313 HisAlaPhePheArgLeuPheGluGluProLysProLysProLysProValLeuPheSer 332
DB 1001 CATGCTTTCTTCAGGCTTTCGGAAGAGCCAAATCAAGCCAAACCTCTCTGCGAGC 1060
QY 333 ArgGlySerSerPheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLys 352
DB 1061 AAGGTTCCAGTTTCCGCTATAGTGAGAACCCAAAGGCAACTTTGGAAATAGGGAGA 1120
QY 353 GlnGlyLysIleValLysValGlnPheGluArgLysHis-----SerLysIleHisSer 370
DB 1121 AAAGGAGGCTGAAGAGCTTCCATTTGAAGAAACATTCACCATCTCAGTACATGAA 1180
QY 371 IleArgSerLeuAlaSerGlnProThrGluLeuAsnSerGluValLeuGlnHisSerGln 390
DB 1181 ---CGACAGTGCAGGTCTCTCAACA---GACCTCTCTGTGATGTCAAAACAA---GTG 1231
QY 391 GlnSerThrSerLeuThrPheGlyGluGly-----Ala 401
DB 1232 GAAGATTGTAGACTAGCATATAGTGAGTGTACTACCAAAATGGAATGAGTGCAGCGCA 1291
QY 402 GluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGlyGlu 421
DB 1292 TCTGAGCGACAGTCTGAGAGAGAGAGAAATCTGTGATGTGAGGTGACATTTTGCAACT 1351
QY 422 ProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAlaAsp 441
DB 1352 GAGCTGGAGCATTCACAA-----CCAGAGCGGATCCCATGTGATCATCAGTCCCAA 1405
QY 442 GlyAlaAlaSerAlaProThrGluGluGluValValLysAspArgThrGlnGln 461
DB 1406 AGCAGTTCTCTTCCCTTT-----ATTATATGAGCCCTGTCTTTAAC 1450
QY 462 SerLysArgGlnPro-----ProGlnPro----- 469
DB 1451 ACTGAGCCCATCTTAACCTCGATCCCAAGACATTTTTCAGAGAGAGTCTTAAGC 1510
QY 470 -----SerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSer----- 484
DB 1511 TCCTTCAAAACMACTGTAAATTTCTGTGTAATACATGACATATATTTCTGGGCTGACA 1570

```

```

QY 485 -----Val 485
DB 1571 AGCAAAAGTGGTCCAGAAAGACGTAACCTTACAGGATGTGCCCTATATCTTGTACA 1630
QY 486 AsnSerGlnGlyValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAsp 505
DB 1631 GGTGACAGGTGGTATATATGCTCCCAAGGTCTTTTATATGAGACAAAGCCACCCAG 1690
QY 506 ThrLysGlnAlaSerProLeuIleSerProLeuLeuAsnArgAlaCysProArgThr 525
DB 1691 GTGCCAGATGTGCCCA----- 1708
QY 526 AspArgGluAspGluGlyArgArgLysArgPheProThrAspLysValTyrPheIleAla 545
DB 1709 -----ATTAGA 1714
QY 546 LysGluValSerThrThrGluArgThrTyrLeuLysAspLeuGluValIleThrSerTrp 565
DB 1715 GCAGAGAAAGACAAAGATCCACATAGCTATGTA----- 1747
QY 566 PheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeuLysSerLeuIlePhe 585
DB 1748 -----GAGCCCATCGCAATGAAGCCAGCTGAAGAGAC----- 1780
QY 586 ProAsnPheGluProLeuHisLysPheHisThrAsnPheLeuLysGluIleGluArg 605
DB 1781 CCAGAAATATCAGAAATGAAGAGCTTTCAGCAAGAC---CTGCAAGTATCTCAAGAAAGCT 1837
QY 606 LeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIle-----ArgAspTyrGlnArg 622
DB 1838 ATAGCCAGCACTAGCGGTAGAGACATCATATAGTGTAGTAAGAGAAAGCAAAAT 1897
QY 623 IleGlyAspValMetLeuLysAsnIleGln 632
DB 1898 TTGGAGATGCATTTGTATGTATCAATTCAA 1927

RESULT 11
AA527016
ID AA527016 standard; cDNA; 1718 bp.
AC AA527016;
DT 07-NOV-2001 (first entry)
XX
DE cDNA encoding novel signal transduction pathway protein, Seq ID 51.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN MO200154733-Al.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US001312.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.

```

PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216860P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226681P.
PR 23-AUG-2000; 2000US-0227009P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 14-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251889P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM,
XX
XX WPI; 2001-465460/50.
DR P-PSDB; AAU17099.
XX
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
PT diagnosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders.
XX
XX
XX Claim 1; SEQ ID NO 51; 880bp; English.
XX
XX The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune

CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative disorders
CC (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.
CC Alzheimer's disease, Parkinson's disease), chromosomal abnormalities
CC (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.
CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),
CC respiratory disorders, dermatological disorders, in wound healing,
CC epithelial cell proliferation, endocrine disorders (e.g. Addison's
CC disease), reproductive system disorders, gastrointestinal disorder
CC (inflammatory disorders), liver disorders (cirrhosis), as stimulators of
CC B-cell responsiveness to pathogens, activators of T-cells, to induce
CC higher affinity antibodies, and as a means to induce tumour proliferation
CC in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-
CC AAS27550 represent novel signal transduction pathway protein coding
CC sequences and PCR primers of the invention
XX

Alignment Scores:
Pred. No.: 1.07e-58 Length: 1718
Score: 1006.50 Matches: 189
Percent Similarity: 78.7% Conservative: 47
Best Local Similarity: 63.0% Mismatches: 57
Query Match: 18.4% Indels: 7
DB: 4 Gaps: 3

US-09-555-342b-2 (1-1045) x AAS27016 (1-1718)

Qy 738 GlnlyLeuHISgluLeuLysAspLeuileGlyIleAspAenLeuValProGly 757
Db 13 CAGAACTACAGGAGCTGACGCGGACCTGGGCGATAGAACTCATCTGCTCGGC 72
Qy 758 ArgGluPheIleArgLeuGlySerLeuSerIlyLeuSerGlyGlyLeuGlnArg 777
Db 73 AGGAGATTATCCGTGAGGGCTGCTTCAAGCTCAAGAGAGGGCTGCACAGAGG 132
Qy 778 MetPhePheLeuPheAspValLeuLeuThrSerArgGlyLeuThrAlaSerAsn 797
Db 133 ATGTTTTCGTTCTCAGATATGTGCTTACCAAGCAAGAGGAGGCTGCACAGAGG 192
Qy 798 GlnPheLeuValHISgluLeuProLeuTyrglyMetThrIleGluGluSerGluAsp 817
Db 193 CACTTCGGGATCCGGGCTCTTCCCTCCAAAGCATGCTGGGAGAAAGATGATAC 252
Qy 818 GluTrpGlyValProHISGlyLeuThrLeuArgGlyGlnArgGlnSerIleIleValAla 837
Db 253 GAGTGTCTGTTCACACTGTTTCAACGCGGCTCAAGAAACATGCTGTGGCA 312
Qy 838 AlaSerSerArgSerGluMetGluLysTrpValGluAspIleGlnMetAlaIleAspLeu 857
Db 313 GCCAGCACTCGGCTGAGAAAGAAAGATGCTGACCTGAACTCCGCGATCCAGCA 372
Qy 858 AlagluYsserSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
Db 373 GCCAAGAGTGGCGGTGACAGGCCCCCTGCCTGCACAGCCGCACTGTGTGACTGCTCC 432
Qy 874 ProAspAnlySerProAspGluAlaThrAlaIAspGlnGluSerGluAspAspLeu 893
Db 433 CCC-----AGATCCCCCAAGAG---GTATCTCTGAGGACGAGTCAAGAGATGATGCT 483
Qy 894 SerIlaSerArgThrSerLeuGluArgGlnAlaProHISargGlyAenThrMetValHIS 913
Db 484 CGGGGTCTCGCACTCCCTGAGAGGGAGTGCAGCAGCCGGGCCAACACACCAATATCAC 543
Qy 914 ValCystrPhtsArgAenThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
Db 544 GTGTGCTGTGATCCGGAACACAGCGTGTCCAGGGCAGAACACCATGACGCTGTGAAAC 603
Qy 934 GlnLeuSerGlyAenLeuLeuArgLysPheIleAspAnSerIleGlyTrpGlnLysLeuTrp 953
Db 604 CAGCTTCAGGATATCTGTCTAAAGAAAGTTCAAAAAGTCAATGCTGGCAGAACTCTGG 663

Qy 954 ValValPheThrAspPheCysLeuPhePheTyrlsSerHISgluAspAnHISProLeu 973
Db 664 GTGTCTTTTACCACTTCTGTGTCTTCTTCTTCAAAATCTATCGATGACTTACCTG 723
Qy 974 AlaSerLeuProLeuLeuGlyTyrglySerLeuThrIleProSerGluSerGluAsnIleGln 993
Db 724 GCCAGCTCCCGGCTGCTGGGTACAGGAGTGCATCCCGAGGAGCGCATGCGATAC 783
Qy 994 LysAspTyrrValPheLysLeuHISPheLysSerHISValTyrrPheAlaGluSer 1013
Db 784 AAAGACTATGTTTCAAGCTCCAGTTCAAATCCACGCTTACTTCCGGGCTGAGAGC 843
Qy 1014 GluTyrrThrPheGluArgTrpMetGluValIleArgSerIlaThrSerSerAlaSerArg 1033
Db 844 AAGTACATTTGAAAGGTGAGTGAAGTATCCAGGAGGCGACGACTCAGCCGGAGG 903
RESULT 12
ADB93194
ID ADB93194 standard; cDNA, 1718 BP.
XX AC
XX ADB93194;
XX 04-DEC-2003 (first entry)
XX DE
XX Human cDNA encoding a novel protein #41.
XX 86; gene; human; autoimmune disease; Parkinson's disease; slllcosis;
XX gastrointestinal disease; atherosclerosis; haemophilia; thrombocytopenia;
XX immunosuppressive agent; adjuvant; enhance immune response;
XX higher affinity antibody induction;
XX increased serum immunoglobulin concentration.
OS Homo sapiens.
PN US2002168711-A1.
XX PD 14-NOV-2002.
XX PF 17-JAN-2001; 2001US-00764868.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 26-JUN-2000; 2000US-0214868P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225278P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234233P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236337P.

PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251865P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 PI Rosen CA, Ruben SM, Barash SC;
 XX WPI; 2003-719985/68.
 DR P-PSDB; ADB93807.
 XX
 PT New isolated polypeptide useful for diagnosing and treating
 PT immunosuppressive conditions such as autoimmune disease and Parkinson's
 PT disease.
 XX
 PS Claim 3; SEQ ID NO 51; 345bp; English.
 XX
 CC The invention relates to an isolated polypeptide. The polypeptide is
 CC useful for diagnosing a pathological condition or a susceptibility to a
 CC pathological condition in a subject, by determining the presence or
 CC amount of expression of the polypeptide in a biological sample and
 CC diagnosing a pathological condition or a susceptibility to a pathological
 CC condition based on the presence or amount of expression of the
 CC polypeptide. The polypeptide is also useful for identifying a binding
 CC partner to the polypeptide, which involves contacting the polypeptide
 CC with a binding partner and determining whether the binding partner
 CC effects an activity of the polypeptide. The polypeptide or the nucleic
 CC acid encoding the polypeptide is useful for preventing, treating, or
 CC ameliorating a medical condition, which involves administering the
 CC polypeptide or the nucleic acid to a mammalian subject. The nucleic acid
 CC is useful for diagnosing a pathological condition or a susceptibility to
 CC a pathological condition in a subject, which involves determining the
 CC presence or absence of a mutation in the nucleic acid, and diagnosing a
 CC pathological condition or susceptibility to a pathological condition
 CC based on the presence or absence of the mutation. The polypeptide, the
 CC nucleic acid and an antibody to the polypeptide are useful for treating
 CC autoimmune disease, Parkinson's disease, sickle cell disease, galectinosis,
 CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,
 CC the nucleic acid and the antibody are useful as immunosuppressive agents,
 CC as adjuvants to enhance immune responses, and as agents to induce higher
 CC affinity antibodies and increase serum immunoglobulin concentrations. The
 CC present sequence represents cDNA encoding a novel human protein. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format direct from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20020168711.
 CC
 XX
 SO Sequence 1718 BP; 428 A; 500 C; 476 G; 314 T; 0 U; 0 Other;

Alignment Scores:
 Pred. NO.: 1.07e-58 Length: 1718
 Score: 1006.50 Matches: 189
 Percent Similarity: 78.7% Conservative: 47
 Best Local Similarity: 63.0% Mismatches: 57
 Query Match: 18.4% Indels: 7
 Gaps: 3

US-09-555-342b-2 (1-1045) x ADB93194 (1-1718)
 QY 738 GlnLysLeuHISgluLeuLysLysAspLeuIleGlyIleAspAsnLeuValProGly 757
 Db 13 CAGAGCTAACCGAGCTGCGAGCGGAGCTGTGGGCAATGAGAACTCATGTCTCTGGC 72
 QY 758 ArgGluPheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnArg 777
 Db 73 AGGAGATTATCATCGTAGGGCTGCTTCACAAAGCTCACCAAGAGAGGCGTCGACAGAGG 132
 QY 778 MetPhePheLeuPheAsnAspValLeuLeuTyrrThSerArgGlyLeuThrAlaSerAsn 797
 Db 133 AGTTTTCCTGTTCTTCACATATGTTGTCGTACACAAAGAAAGAGTTCAGGGACCGC 192
 QY 798 GlnPheValHISgluLeuProLeuTyrrGlyMetThrIleGluGlnSerGlyAsp 817
 Db 193 CACTTCGGATCCGGGCTCTCTCTCCCTCCCAAGGATCTGTGTGAAGAAATGATTAAC 252
 QY 818 GluTrpGlyValProHISCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAla 837
 Db 253 GAGTGTCTGTTCCACACTGTTTCACCATCTACCGCGCTCAGAAACAAATCGTGTGGCA 312
 QY 838 AlaSerSerArgSerGlyMetGlyLysTrpValGluAspIleGlnMetAlaIleAspLeu 857
 Db 313 GCCAGCACTCGGCTGAGAAAGAGTGAATGATCTGGACCTGACCTCGCATTCAGCA 372
 QY 858 AlaGluLysSerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
 Db 373 GCCAAGATGGCGGTACACGCGCTTCGATCCAGCGCGGCTGCTGTGCTCTCTCC 432
 QY 874 ProAspAsnLysSerProAspGluAlaThrAlaAlaAspGlnGlnSerGlyAspAspLeu 893
 Db 433 CCC-----AGATCCCCCAAGCAG--GATCTCTGAGACAGAGATCAGAAATGATGCT 483
 QY 894 SerAlaSerArgThrSerLeuGlnArgGlnAlaProHISArgGlyAsnThrMetValHIS 913
 Db 484 CCGGAGTCCCGAGCTCTCCCTGGAGGGGATGGCCAGCACCGGCGCAACACCAATGAC 543
 QY 914 ValCysTrpHISArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
 Db 544 GTGTGCTGTGATCCGGAACACACGCGTTCAGGGCAGACACACAGTCCAGTCTCGAAGAC 603
 QY 934 GlnLeuSerGlyAsnLeuLeuArgLysPheLysAsnSerAsnGlyTrpGlnLysLeuTrp 953
 Db 604 CAGCTTTCAGATATCTGCTAAGAAAGTTCAAAACAGTCATGCTGCGCAAGGCTCTGG 663
 QY 954 ValValPheThrAsnPheCysLeuPhePheTyrrLysSerHISGlnAsnHISProLeu 973
 Db 664 GTGCTTTTACCAACTTGTGTTGTTCTTCAAAAACATCATTGAGATGACTACCACTG 723
 QY 974 AlaSerLeuProLeuLeuGlyTrpSerLeuThrIleProSerGlnSerGluAsnIleGln 993
 Db 724 GCCAGCTCCCGCTGTGGGCTTACAGCGTGCAGATCCCAAGAGGCCATGGCATACAC 783
 QY 994 LysAspTrpValPheLysLeuHISPheLysSerHISValTyrrTrpPheArgAlaGlnSer 1013
 Db 784 AAAGACTAATGTTTCAAGCTCCCAATTCCAATCCCAAGTCTACTCTCTCCGGCTGAGAGC 843
 QY 1014 GluTyrrThrPheGlnArgTrpMetGluValIleArgSerAlaThrSerSerAlaSerArg 1033
 Db 844 AAGTACACATTTGAAAGGTGATGAGGTGATTCAGAGGGGCGCAGAGCTCAGCGGAGG 903
 RESULT 13
 ACD92278/c
 ID ACD92278 standard; cDNA; 582 BP.
 XX
 AC ACD92278;
 XX
 DT 23-SEP-2003 (first entry)
 XX
 DE Human colon cancer cell expressed cDNA #690.
 KW Open reading frame detection; genome sequencing; colon cancer;

KM breast cancer; population genome analysis; genetic shift; cancer;
 KM antibiotic resistance; antibiotic non-tolerance; congenital disease;
 KM agriculture; food crop genome; resistance gene; retrovirus;
 KM influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
 gene; ss.
 XX Homo sapiens.
 XX US2002155438-A1.
 XX
 XX 24-OCT-2002.
 XX
 XX 27-SEP-1999; 99US-00406117.
 XX
 XX 20-NOV-1998; 98US-00196716.
 XX
 XX (SIMP/) SIMPSON A J G.
 XX (NETO/) NETO E D.
 XX (BREN/) BRENTANI R R.
 XX
 XX Simpeon AJG, Neto ED, Brentani RR;
 DR WPI; 2003-182626/18.
 XX
 PT Determining open reading frames of genome of an organism e.g. a human
 PT suffering from cancer involves use of single oligonucleotide primer at
 PT low stringency for preparing single-stranded cDNA from mRNA of
 PT individual.
 XX
 XX Example 9; Page 105; 959pp; English.
 XX
 CC The invention describes a method of determining open reading frames in
 CC the genome of organism, comprising contacting mRNA from cell of organism
 CC with a single oligonucleotide primer (I) at low stringency, preparing
 CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
 CC cDNA, sequencing the product, and repeating the contacting, preparing
 CC and amplifying steps with different primers and sequencing resulting
 CC nucleic acids. The method is useful for: determining that a known
 CC nucleotide sequence from a genome of an organism corresponds to a
 CC nucleotide sequence of an open reading frame; for preparing a contig,
 CC nucleic acid molecule from a genome of an organism; and for sequencing
 CC all or part of a genome of an organism. mRNA is obtained from mammalian
 CC or human cell which is associated with a pathological condition e.g. a
 CC colon cancer or breast cancer cell. The method is useful for analyses of
 CC populations of subjects and can be used to carry out genetic analyses of
 CC large or small populations. Further, it can be used to study living
 CC systems to determine if, e.g. there have been genetic shifts which render
 CC an individual or population more or less likely to be afflicted with
 CC diseases such as cancer, to determine antibiotic resistance or non-
 CC tolerance, and so forth. The method can also be used in the study of
 CC congenital diseases, and the risk of affliction to a foetus, as well as
 CC the study of whether the conditions are likely to be passed to offspring
 CC through ova or sperm. The analyses for pathological conditions can be
 CC carried out in all animals, plants, birds, fish, etc. Using this method,
 CC in the area of agriculture, for example the genomes of food crops can be
 CC studied to determine if resistance genes are present, defects in plant
 CC genomes can also be studied in this way. Similarly, the method permits
 CC determination of the pathogens which integrate into the genome, such as
 CC retroviruses and other integrating viruses such as influenza virus, have
 CC undergone shifts or mutations, which may require different approaches to
 CC therapy. This method is also applied to eukaryotic pathogens, such as
 CC trypanosomes, different types of Plasmodium, etc. The method essentially
 CC eliminates sequencing of non-coding portions. This sequence represents a
 CC polynucleotide isolated from human colon cancer cell cDNA library
 XX
 XX Sequence 582 BP; 109 A; 147 C; 185 G; 140 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 2.01e-56
 Score: 964.00
 Percent Similarity: 98.4%
 Best Local Similarity: 97.9%
 Query Match: 17.6%
 Matches: 582
 Conservative: 183
 Mismatches: 3
 Indels: 0

DB: 10 Gaps: 0
 US-09-555-342B-2 (1-1045) x ACD92278 (1-582)
 QY 859 GluYSerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLysSer 878
 Db 579 GAAAGAAACAGACAGACCCGCCCTTGAAGTTCCTGGCCAGACGCCCTCGACAAACAGTCC 520
 QY 879 ProAspGluAlaThrAlaAlaAspGlnGluSerGluAspAspLeuSerAlaSerArgThr 898
 Db 519 CCGTATAGACCAACCGCGGCTGACACAGAGTCAAGATGACCTGACCGCTTCGACCA 460
 QY 899 SerLeuGluArgGlnAlaProHlaArgGlyAsnThrMetValHisValCysTrpHisArg 918
 Db 459 TCGCTGAGCGCCAGCGCCCGCCAGCCGCGCAACCAATGTGTGACGTCTGACACCGC 400
 QY 919 AsnThrSerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerGlyAsn 938
 Db 399 AACACACAGCGTCTCATGTGAGTCACTCAGCATCGCAGTGGAGATCAGTTGTCTGAAAC 340
 QY 939 LeuLeuArgLysPheLysAsnSerAsnGlyTrpGlnLysLeuTrpValAlaPheThrAsn 958
 Db 339 CTGCTGAGGAAATTCAAAAACAGCAACGGGTGGCAAGCTGTGGTGTTCACAAAC 280
 QY 959 PheCysLeuPhePheTyrlYsSerHisGlnAspAsnHisProLeuAlaSerLeuProLeu 978
 Db 279 TTCGCTGCTTCTTCTTCAAAATCACACAGACCAATATCCTTGGCAGCTGCTCTG 220
 QY 979 LeuGlyTyrlSerLeuThrIleProSerGluSerGluAsnIleGlnLysAspTyrlValPhe 998
 Db 219 CTGGCTACTCGCTCACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTAGCTGTTCC 160
 QY 999 LysLeuHisPheLysSerHisValTyrlTyrlPheArgAlaGluSerGlyTrpPheGlu 1018
 Db 159 AAGCTGACCTTCAAGTCCAGCTTACTTCTTCAAGGCGGAAACGAGTACACCTTGAA 100
 QY 1019 ArgTrpMetGluValIleAspSerAlaThrSerSerAlaSerArgProHisValLeuSer 1038
 Db 99 AGGTGATGAGATGATCCGACGTGCCACCAAGCTTCTGCGGACCCAGTGTTAGAC 40
 QY 1039 HisLysGluSerLeuValTyrl 1045
 Db 39 CACAAAGAGTCTTGTGTAT 19
 RESULT 14
 ACD94832/c
 ID ACD94832 standard; cDNA; 582 BP.
 XX
 AC ACD94832;
 XX
 DT 23-SEP-2003 (first entry)
 XX
 DE Human colon cancer cell expressed cDNA #3244.
 XX
 KM Open reading frame detection; genome sequencing; colon cancer;
 KM breast cancer; population genome analysis; genetic shift; cancer;
 KM antibiotic resistance; antibiotic non-tolerance; congenital disease;
 KM agriculture; food crop genome; resistance gene; retrovirus;
 KM influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
 gene; ss.
 XX
 XX Homo sapiens.
 XX US2002155438-A1.
 XX
 XX 24-OCT-2002.
 XX
 XX 27-SEP-1999; 99US-00406117.
 XX
 XX 20-NOV-1998; 98US-00196716.
 XX
 XX (SIMP/) SIMPSON A J G.
 XX (NETO/) NETO E D.

XX	(BREN/) BRENTANI R. R.
XI	Simpson AJG, Neto ED, Brentani RR;
XX	WPI; 2003-182626/18.
DR	
XX	Determining open reading frames of genome of an organism e.g. a human
PT	suffering from cancer involves use of single oligonucleotide primer at
PT	low stringency for preparing single-stranded cDNA from mRNA of
PT	individual.
PS	
XX	Example 9; Page 477; 959pp; English.
CC	The invention describes a method of determining open reading frames in
CC	the genome of organism, comprising contacting mRNA from cell of organism
CC	with a single oligonucleotide primer (I) at low stringency, preparing
CC	single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
CC	cDNA, sequencing the product, and repeating the contacting, preparing
CC	and amplifying steps with different primers and sequencing resulting
CC	nucleic acids. The method is useful for: determining that a known
CC	nucleotide sequence from a genome of an organism corresponds to a
CC	nucleotide sequence of an open reading frame; for preparing a contig,
CC	nucleic acid molecule from a genome of an organism; and for sequencing
CC	all or part of a genome of an organism. mRNA is obtained from mammalian
CC	or human cell which is associated with a pathological condition e.g. a
CC	colon cancer or breast cancer cell. The method is useful for analyses of
CC	populations of subjects and can be used to carry out genetic analyses of
CC	large or small populations. further, it can be used to study living
CC	systems to determine if, e.g. there have been genetic shifts which render
CC	an individual or population more or less likely to be afflicted with
CC	diseases such as cancer, to determine antibiotic resistance or non-
CC	tolerance, and so forth. The method can also be used in the study of
CC	congenital diseases, and the risk of affliction to a foetus, as well as
CC	the study of whether the conditions are likely to be passed to offspring
CC	through ova or sperm. The analyses for pathological conditions can be
CC	carried out in all animals, plants, birds, fish, etc. Using this method,
CC	in the area of agriculture, for example the genomes of food crops can be
CC	studied to determine if resistance genes are present, defects in plant
CC	genomes can also be studied in this way. Similarly, the method permits
CC	determination of the pathogens which integrate into the genome, such as
CC	retroviruses and other integrating viruses such as influenza virus, have
CC	undergone shifts or mutations, which may require different approaches to
CC	therapy. This method is also applied to eukaryotic pathogens, such as
CC	trypanosomes, different types of Plasmodium, etc. The method essentially
CC	eliminates sequencing of non-coding portions. This sequence represents a
CC	polynucleotide isolated from human colon cancer cell cDNA library
XX	
SQ	Sequence 582 BP; 109 A; 147 C; 185 G; 140 T; 0 U; 1 Other;
Alignment Scores:	
Pred. No.:	2, 01e-56 Length: 582
Score:	964.00 Matches: 183
Percent Similarity:	98.4% Conservative: 1
Best Local Similarity:	97.9% Mismatches: 3
Query Match:	17.6% Indels: 0
DB:	10 Gaps: 0
US-09-555-342B-2 (1-1045) x ACD94832 (1-582)	
OY	859 GlutylaseSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLysSer 878
OY	::::: 579 GAAGAAGACAGCAGCCCCCGCTGTAGATTCCTGGCCAGCAGGCCCTCGAACAACATCC 520
Db	879 ProAbpGluAlaThrAlaAlaAbpGlnGluSerGluAbpAbpLeuSerAlaSerArgThr 899
Db	519 CTGATGAACCCACCGCGGTGCACGAGAGTCAAGMATTGACTTGAGGCTCTCCGCA 460
OY	899 SerLeuGluArgGlnAlaProHisAlaArgGlyValThrMetValHisValCysTrpHisArg 918
Db	459 TCCTGAGAGCGCACAGGCCCGCCGACCGCGGCAACAAATGATGACATGTCTGCGACACCC 4000
OY	919 AsnThrSerValSerMetValAbpPheSerIleAlaValGluAbpGlnLeuSerGlyAsn 938

Db	399	AACACACGACGCTTCCTCCAGTGGAGCTTACAGATGCGACGTGAGAAATCAGTTGCTGGAAAC	340
QY	939	LeuLeuAArgLySpPheLySbaenSerAenGlyTTPGInLyLeuTPValValPheThrAsn	958
Db	339	CTGCTGAGGAAATTCAAAACAGACACGGGTGGAGAAAGCTGGGGGTGTTCACAAAC	280
QY	959	PheCyAsuPhePhePheTyrLySserSerHisGlaAspAsnHisProLeuHisSerLeuProLeu	978
Db	279	TTTCGCTGCTTCTTCTCAAAATCACACACAGACAAATATATCCCTTGGCCAGCTGCTCTG	220
QY	979	LeuGlyTyrSerLeuThrIleProSerGlySerGlyuAsnIleGlnLySbaPyrValPhe	998
Db	219	CTCGGCTACTGCTGCACCAATCCCTCTGAGTCCGAGAAATCCAGAAAGACTGCTGTTTC	160
QY	999	LyLeuLeuHisPheLySserSerHisGlyTyrTyrPheArgHisGlaLySerGlyTyrThrPheGlu	1018
Db	159	AAGCTGCACTTCAAATGCCACGCTTACTTTCAGGGCGGAAAGCCGATACAGTTGAA	100
QY	1019	ArgTyrPheGlyuValIleArgSerLarThrSerSerHisSerArgProHisValLeuSer	1038
Db	99	AGGTGGATGGAAAGGATCCGACAGTGCACACAGCTCTGCTCGGAGACCCAGTGTAGC	40
QY	1039	HisLySgluSerLeuValTyr	1045
Db	39	CACAAAGAGTCTTGTGTAT	19
RESULT 15			
ABL03867	ID	ABL03867 standard; cDNA; 1747 BP.	
XX	AC	ABL03867;	
XX	DT	26-MAR-2002 (first entry)	
XX	DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 6083.	
XX	KM	Drosophila; developmental biology; cell signalling; insecticide;	
XX	KM	pharmaceutical; gene; ss.	
XX	OS	Drosophila melanogaster.	
PN	WO200171042-A2.		
PD	27-SEP-2001.		
XX	PF	23-MAR-2001; 2001WO-US009231.	
XX	PR	23-MAR-2000; 2000US-0191637P.	
PR	11-JUL-2000; 2000US-00614150.		
XX	PA	(PEKE) PE CORP NY.	
XX	PI	Venter JC, Adams M, Li PWD, Myers EW;	
XX	DR	WPI; 2001-656860/75.	
XX	DR	P-PSDB; ABB59764.	
PT	FT	New isolated nucleic acid detection reagent for detecting 1000 or more	
XX	PT	genes from Drosophila and for elucidating cell signaling and cell-cell	
XX	XX	interactions.	
PS	Claim 1; SEQ ID NO 6083; 21bp + Sequence Listing; English.		
CC	XX	The invention relates to an isolated nucleic acid detection reagent	
CC	CC	capable of detecting 1000 or more genes from Drosophila. The invention is	
CC	CC	useful in developmental biology and in elucidating cell signaling and	
CC	CC	cell-cell interactions in higher eukaryotes for the development of	
CC	CC	insecticides, therapeutics and pharmaceutical drugs. The invention	
CC	CC	discloses genomic DNA sequences (AB116176-AB130511), expressed DNA	
CC	CC	sequences (AB101840-AB116175) and the encoded proteins (ABB57737-	
CC	CC	ABB572072). The sequence data for this patent did not form part of the	
CC	CC	printed Specification, but was obtained in electronic format directly	
CC	CC	from WIPO at ftp.wipo.int/pub/published/pct_sequences	

XX Sequence 1747 BP: 431 A; 492 C; 461 G; 363 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,26e-52 Length: 1747
 Score: 913.50 Matches: 217
 Percent Similarity: 54.0% Conserved: 110
 Best Local Similarity: 35.8% Mismatches: 207
 Query Match: 16.7% Indels: 73
 DB: 4 Gaps: 14
 US-09-555-342B-2 (1-1045) x ABL03867 (1-1747)
 QY 423 GlyserrHsProSerProAlaProArgSerProAlaGlyAenlySGlnAlaAspGly 442
 Db 87 GGATCGGATTAAGCTTTCCTCACTATCCCGATCTGAGCGAACA----- 131
 QY 443 AlaAlaSerAlaProThrGlnGluGluGluValValLysAspArgThrGlnGlnSer 462
 Db 132 -----TATGACGTAAACAGCGAAGAAATCCAGACGCT 164
 QY 463 LysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeuSerGlu 482
 Db 165 AAGAGACAG-----GAGCTAGCCACAGCGCTGCCACCGCTCG----- 203
 QY 483 LeuSerValAsnSerGlnGlyValAlaProAlaAsn---ValThrLeuSerProAsn 501
 Db 204 -----CACCGAATGGAGACGGAACCGAAATGGGCATACCTTAGC----- 245
 QY 502 LeuSerProAspThrLysGlnAlaSerProLeuLeuSerProLeuLeuAsnAspGlnAla 521
 Db 245 ----- 245
 QY 522 CysProArgThrAspAspGluAspGluValArgArgLysArgPheProThrAspLysAla 541
 Db 246 ---ACACAAACAGATATTGAGCGCAAGTGAAGAAAGCGGAATGGCCACCGAGCCGAC 302
 QY 542 TyrPheIleAlaLysGluValSerThrThrGlnArgThrTyrLeuLysAspLeuGluVal 561
 Db 303 TACTTTTGGCCAAAGAGAGCTGTATGACGGACGTACGTAACAAAGAGATCTAGAGTGTG 362
 QY 562 IleThrSerTrpPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeuLys 581
 Db 363 CTGAACACACCTTCCACAGCTGTAGTCTCGGGAGTGTG-----GACACCTGTGAG 416
 QY 582 SerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAsnPheLeuLysGlu 601
 Db 417 CCGCTG---TTCGAGCTGCTGACCTCCCTGGCCGACACCAATCTCTCTGCGGAC 473
 QY 602 IleGluGlnAlaGluAlaLeuTrpGluGlyArgSerAsnAlaGlnIleArgAspTyrGln 621
 Db 474 ATCCAAACACCCCATGCTGTCAGATGGAGGAGGAGAGAGCGC-----CACGAAGCCGAC 524
 QY 622 ArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLysAlaAlaHisLeu 641
 Db 525 CGCATCGGAGACGTCATGATAGACATGGCCGCTGTGCCCCATCTACGATGAGTACGTG 584
 QY 642 TrpLysHisSerGlnAlaLeuGluAlaLeuGlnGlyLysLysSerSerArgArgLeu 661
 Db 585 CAGACGACCTGACATCTTGCACTGTATGAACAGCATGTACGAAGGAGGAGAACGCTTC 644
 QY 662 GluAsnPheCysAspAspPheGluLeuGlnLysValCysTyrLeuProLeuAsnThrPhe 681
 Db 645 CGTCAGAGGTGACAGAGGTTTGACGACAAAGGTTTGCTATCTAACCATGAGGGAATTT 704
 QY 682 LeuLeuArgProLysHisArgLeuMetHisTyrLysGlnValIleGluArgLeuCysLys 701
 Db 705 CTACTGAAGCCCTCCACCGCTGTGCACTAACCACTGATCTTAGAGCGGCTTGGCAC 764
 QY 702 HisHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIleThr 721
 Db 765 TACTATGGGAGAGGAGCATATGCACTATGCGATGCTATGCGCGTACCACTTGCTGCTT 824

QY 722 GluMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPheGlnLysLeuHis 741
 Db 825 CGCAGCACCAAGGATATAGTCCAGACTCCGACCTGCAACTTTCGAGAGCTGTGT 884
 QY 742 GluLeuLysLysAspLeuIleGlyIleAspAsnLeuValValProGlyArgGluPheIle 761
 Db 885 GAATCGAAGCGCAC---ATCAACTTGACAGAGTGTGTACAGCCCTTCCCGCTTATC 941
 QY 762 ArgLeuLysSerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArgMetPhePheLeu 781
 Db 942 CGCCAGGATGCTTCTGAAGACCTCCAGCGCGGTGTGACAGAGAGATCTTGTG 1001
 QY 782 PheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsnGlnPheLysVal 801
 Db 1002 TTCTCCGACCTGCTGCTCTTACGTTACGAATGCC---CCGCTGACCAAGCTTCCGAT 1058
 QY 802 HisGlyLeuLeuProLeuTyrGlyMetThrIleGluGlnSerGluAspGluTrpGlyVal 821
 Db 1059 TTAGGCCACGTTCTGTCGCTTCACTACACCGAAGCCCGAGCAC----- 1106
 QY 822 ProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAlaSer----- 839
 Db 1107 ---AACACTTCTCCATCTTCGGCGACAGTGTGCCATCAGTGAACAAGATGCACC 1163
 QY 840 -----SerArgSerGluMetGluLysTrpValGluLysPheIleGlnMetAlaIleAspLeu 857
 Db 1164 CAATCCCAAAATCTTCTTCCACAAAGCACCCAGTAATGTATTCCTACCTTATTTTCA 1223
 QY 858 AlaGluLysSerSerProAlaProGluPheLeuAlaSerSerProProAsnLys 877
 Db 1224 GGCTCTTC-GAAGAGGCGCTTGACCT-----ATTGCGTGTGAGCAACGGGAACAACAG 1276
 QY 878 SerProAspGluAlaThrAlaAlaAspGlnGlnSerLys-----AspAspLeuSer 894
 Db 1277 CAGCTTAAACAGCAGTGTGAACGGCGCGGTCCGTGACTTACGACGACAGAGAACTTGA 1336
 QY 895 AlaSerArgThrSerLeuGluArgGlnAla-ProHisArgGlyAsnThrMetValHisVal 914
 Db 1337 GCTGCACCAACAGCAGCAAAACAGAGCAGCCATCCCAAGCAATACGCTGCGACGT 1396
 QY 914 LysTrpHisArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsnGln 934
 Db 1397 CTGTGCGATCGTGGCCACCGTACGAGTGGGCGATCACCTGATACCCCGGACACCA 1456
 QY 934 PheSerGlyAsnLeuLeuArgLysPheLysAsnSerAsnGlyTrpGlnLysLeuTrpVal 954
 Db 1457 GTTGTCCGTTATCTGTCTGGAAAGTTCAAGAAATAGCTCCGCTGGCAGAACTTGGGT 1516
 QY 954 ValPheThrAsnPheCysLeuPhePheTyrLysSerHisGlnAspAsnHisProLeuAla 974
 Db 1517 GGTGTACAGCTCTTTGTGTCTGACTTTCACAAAGCTTACAGGACAGATTCGACATGGC 1576
 QY 974 AserLeuProLeuLeuGlyTyrSerLeuThrIleProSerGlnSerGluAsnIleGlnLys 994
 Db 1577 TAGCTCCCGCTATATGGTTTACACGATGGGCTCCCTGTCACCAAGATGCCGTTCAAA 1636
 QY 994 AspTrpValPheLysLeuHisPheLysSerHisValTyrTyrPheArgAlaGlnSerGln 1014
 Db 1637 GAACTTGCTCTTCAAGCTTCTTGAAGAACACAGTCTACTTCTTCGCGCGGAAAGTGC 1696
 QY 1014 uTyrThrPheGluArg 1019
 Db 1697 GCACACCTTACAAACAGG 1712

Search completed: May 20, 2006, 03:10:22
 Job time : 1574 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 BioCeleration Ltd.

Om protein - nucleic search, using frame_plus model

Run on: May 20, 2006, 01:45:51 ; Search time 11213 Seconds

(without alignments)
7817.137 Million cell updates/sec

Title: US-09-555-342B-2

Perfect score: 5463

Sequence: 1 MGELQRPRTGSRGAPENS.....SATSSARPHVLSHESLVY 1045

Scoring table:
BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ .p2n.model -DEV=x1h
-Q=/abs/ABSMBR.spool/US09555342/runat_19052006_121457_13216/app_query.fasta_1
-DB=EST -QMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFCR=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blowsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abse07
-USRR=US09555342.0CGN.1.1.6323.0runat_19052006_121457_13216 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1.*
2: gb_est3.*
3: gb_est4.*
4: gb_est5.*
5: gb_est6.*
6: gb_est7.*
7: gb_est8.*
8: gb_est9.*
9: gb_est10.*
10: gb_est11.*
11: gb_est12.*
12: gb_est13.*
13: gb_est14.*
14: gb_est15.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5463	100.0	3138	14	AY406167 Homo sapi
2	5359	98.1	4775	6	CR859100 Pongo pyg
3	5116	93.6	3138	14	AY406168 Pan trogl
4	5054.5	92.5	3147	14	AY406169 Mus muscu

Result No.	Score	Query Match	Length	DB ID	Description
5	2916	53.4	3743	6	AK171713 Mus muscu
6	2896.5	53.0	3874	6	AK050860 Mus muscu
7	2318.5	42.4	3165	14	AY415515 Homo sapi
8	2286	41.8	3168	14	AY415517 Mus muscu
9	1920.5	35.2	3126	14	AY415516 Pan trogl
10	1648	30.2	968	4	BX364606 BX364606
11	1615	29.6	1000	4	BX354921 BX354921
12	1581.5	28.9	1354	10	DV784833 Hw liver
13	1544	28.3	1026	2	BMS51766 AGENCOURT
14	1514	27.7	912	3	BQ881649 AGENCOURT
15	1496	27.4	904	3	BQ931960 AGENCOURT
16	1463.5	26.8	959	3	BUS39080 AGENCOURT
17	1448	26.5	1072	4	BX354920 BX354920
18	1445.5	26.5	915	3	BH846495 AGENCOURT
19	1445.5	26.5	1043	3	BQ072025 AGENCOURT
20	1444	26.4	926	3	BQ706499 AGENCOURT
21	1442	26.4	857	4	BX50128 BX50128
22	1441.5	26.4	966	3	BUI46764 AGENCOURT
23	1418	26.0	889	3	BUI74449 AGENCOURT
24	1395	25.5	951	4	BX364605 BX364605
25	1383	25.3	806	10	DR763338 HESCA_151
26	1382.5	25.3	923	3	BQ950768 AGENCOURT
27	1380	25.3	2161	6	AK034903 Mus muscu
28	1373	25.1	800	9	CX781945 HESCA_10
29	1363.5	25.0	899	4	BX327170 BX327170
30	1357	24.8	896	3	BQ672632 AGENCOURT
31	1355	24.8	1122	7	BE746268 AGENCOURT
32	1344	24.6	800	5	CD351400 UI-M-G10
33	1341.5	24.6	881	3	BUI63030 AGENCOURT
34	1324	24.2	804	2	BG747792 AGENCOURT
35	1321	24.2	774	2	B1760610 603044980
36	1319.5	24.2	930	1	AL566821 AL566821
37	1319	24.1	861	2	BG764061 602737071
38	1310	24.0	804	8	CNS28131 UI-M-HQ0
39	1306	23.9	784	2	BG763918 602736888
40	1299	23.8	773	4	CB519600 UI-M-GH0
41	1294	23.7	2816	6	AK086699 Mus muscu
42	1289	23.6	952	3	BQ676994 AGENCOURT
43	1288.5	23.6	799	9	CX568240 AGENCOURT
44	1287.5	23.6	865	7	BE260677 601154188
45	1282	23.5	957	3	BUS20574 AGENCOURT

ALIGNMENTS

RESULT 1
AY406167
LOCUS
DEFINITION Homo sapiens FARPI gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY406167
VERSION AY406167.1 GI:39762141
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Fertigera,S., Wang,G., Zheng,X.H., White,T.J., Shteyn,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Fertigera,S., Wang,G., Zheng,X.H., White,T.J., Shteyn,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission

JOURNAL

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

COMMENT

This sequence was made by sequencing genomic exons and ordering
them based on alignment

FEATURES

Location/Qualifiers
1..3138

SOURCE

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..3138
/gene="PAPPI"
/locus_tag="HCM2467"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	3138
Score:	5463.00	Matches:	1045
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	14	Gaps:	0

US-09-555-342B-2 (1-1045) x AY406167 (1-3138)

QY 1 MecG1yG1u1egiug1nArqProThrProG1ySer4rgleuG1yAlaProG1uA5nSer 20
Db 1 ATGGAGAAATTAAGACAGAGCGCGACCCAGATCAAGACTGGGGGCCCCGAAATTCG 60
QY 21 G1y1SerThrleuG1uArqG1yG1nLpProBProThrProSerG1yLpLeuVal 40
Db 61 GGGATCGTACCTTGGAACGTGACAGAACGCCGCCCAACCTTCAGGAAACCTCGTG 120
QY 41 Ser1eLys1egiMeLleuAAsPArThrG1nG1uAlaPheG1uValProG1nArqAla 60
Db 121 TCCATCAAAATCCAGATGCTGGATGACACCCAGAGGACTTTGAACTTCCAAAGAGCT 180
QY 61 ProG1yLpValLeuLeuAAsPAlaValCysA5nH1A5eubAnLeuValG1uG1yAsP1y 80
Db 181 CCTGGAGAGGTGCTGATGATGAGTTTGCAACCTCAACCTCGTGAAGGTGATCTAT 240
QY 81 PheG1yLeuG1uPheProAAsPAlaValLys1eThrVal1TrpLeuAAsPLeuLpPro 100
Db 241 TTTGGCTCGAGTTCTGATCACAAAAGATCAAGGTGCTGAGATCTCTTAAACCC 300
QY 101 1LeVal1ySG1n1eArqArqProLp5H1eVal1Val1yLpPheVal1yLpPhe 120
Db 301 ATTGTAAACGATTAAAGGCCAAAGCACTTGTGTTAAAGTTGTGGTAAATCTTT 360
QY 121 ProBProAAsP1eThrG1nLeuG1nG1uLeuThraTrq1yLeuPheAlaLeuG1nVal 140
Db 361 CCGGCTGACCAACACCACTCAAGAAAGTCAAGAGTACCTGTTCCGGCTGACAGTG 420
QY 141 LysG1nAAsPLeuAlaG1nG1yArqLeuThrCysA5nAAsPThrSerAlaAlaLeuLeu11e 160
Db 421 AAGAGAGACTTGGCTCAAGGAGGTTGACGTGAATAGACACAGGAGCTCTCTGATT 480
QY 161 Ser1e1LeVal1G1nSerG1u11eG1yAAsPheAAsP1uAlaLeuAAsPArqG1uH1eLeu 180
Db 481 TCACACATTTGTGCAATCTGAGATTGGGGATTTGATGAAGCTTGGACAGAGACACTTA 540
QY 181 AlAlaLysAAsP1yTr1eProG1nG1nAAsPAlaLeuG1uAAsP1yAlaG1uPheH1e 200
Db 541 GCAAAATAATTAATCATACCTCAAGCAAGACGCACTAGAGGCAAAATCGTGAATTTTCA 600
QY 201 H1AAsnH1e11eG1yG1nThrProAlaG1uSerAAsPheG1nLeuLeuG1u11eAlaArq 220
Db 601 CATTAACACATTTGACAAACCAAGCAAGATCAAGATTTCCAGCTTCAAGATTTGCCGT 660
QY 221 ArqPheG1uMe1yArq1y11eArqLeuH1eProAlaLysAAsPArqG1uG1yThrLys11e 240
Db 661 CGGCTTAAGATGATTAAGATCGGTTGACCCGGCCAAAGACAGAGGAGGACCAAGATC 720
QY 241 A5nLeuAlaValAlaA5nThrG1y1LeuVal1PheG1nG1yPheThrLys11eA5nAla 260

Db 721 AATCTGCGCTTGCACAAACGGGAATTCATGTTTTCAGGGTTTCACTAAGACAAATGCC 780
QY 261 PheA5nThrAlaLysValArqLysLeuSerPheLysArqLysArqPheLeu11eLysLeu 280
Db 781 TTCACTGGAGCCAGAGGTGGAGAGCTGAGCTTCAAGAGAGAGGCTTTCTCATCAAGCTC 840
QY 281 ArqProAAsPAlaA5nSerAlaTrqG1nAAsPThrLeuG1uPheLeuA5eArqAAsP 300
Db 841 CGGCGAGAGCCAAATGTTGATCCAGGATACCTTGAAATTCCTGATGGCCAGTCGGGAT 900
QY 301 PheCylusSerPheTrpLys11eCysValG1uH1eH1eAlaPhePheA5eLeuPheG1u 320
Db 901 TTCTGCAAGCTCTTCGAAAAATCTGTGTGAACATCATGCTCTTGTGAACCTTTTGA 960
QY 321 GluProLysProLysProLysProValLeuPheSerArqG1ySerSerPheArqPheSer 340
Db 961 GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGGTCAATCATTTCCGTTCACT 1020
QY 341 G1yArqThrG1nLpSG1nValLeuAAsP1yVal1ySG1uG1yG1yH1eLysValG1n 360
Db 1021 GGTGGACTCAGAAAGCAGGTTCTCACTATGTTAAAGAGAGACATAAGAGGTGCAG 1080
QY 361 PheG1uArqLysH1eSerLys11eH1eSer11eArqSerLeuAlaSerG1nProThrG1u 380
Db 1081 TTTGAAGAGAGACAGCAAGATTCATTCATCCGAGCCTTCTTCAACGCTTACAGAA 1140
QY 381 LeuA5nSerG1uValLeuG1nSerG1nSerThrSerLeuThrPheG1yG1uG1y 400
Db 1141 CTAAATTCGAAATGCTGAGAGCTTCAGCAAGACACAGCCTTACATTTGAGAAAGT 1200
QY 401 AlAG1uSerProG1yG1nSerCysArqArqG1yLpSG1uProLysVal1SerAlaG1y 420
Db 1201 GCCGAATCTCCAGAGGGGCGAGACTGCGCGGAGAAAGAACGAAAGTTTCCGCGGG 1260
QY 421 GluProG1ySerH1eProSerProAlaProArqArqSerProAlaG1yA5nLpSG1nAla 440
Db 1261 GAGCCGGGGTGGACCCGAGCCTGCGCGAGAGAAAGCCCGGGGTAAACAAGAGCGG 1320
QY 441 AAsP1yAlaAlaSerAlaProThrG1uG1uG1uG1uVal1yLysAAsPArqThrG1n 460
Db 1321 GACGAGAGCGCTCGGCGGCCACGAGAGAGAGAGAGAGTCTTAAAGATAGACCCAG 1380
QY 461 G1nSerLysProG1nProProG1nProSerThrG1ySerLeuThrG1ySerProH1eLeu 480
Db 1381 CAGAGTAAACTCAGCCCGGACGACAGCAAGAGCTCCTGATCGGCACTGCTCACTT 1440
QY 481 SerG1uLeuSerValA5nSerG1nG1yValAlaProAlaA5nVal1ThrLeuSerPro 500
Db 1441 TCCGAGCTGTGTGAACCTGCAAGGGGGAGTGGCCCTGCAACGTGACTGTCTCC 1500
QY 501 A5nLeuSerProAAsPThrLysG1nAlaSerProLeu11eSerProLeuLeuA5nAAsP1n 520
Db 1501 AACCTGAGCCCAACACAGAGAGGCTCTCTTATCAGCCCGCTGCTGAATGACAG 1560
QY 521 AlACysProArqThraAsPArq1uAAsPArqG1yArqArq1yAaTrqPheProThraAsP1y 540
Db 1561 GCTGCGCCCGGACGAGCATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGAATAAA 1620
QY 541 AlArqTrPhe11eAlaLysG1uVal1SerThrThrG1uArqThrTrqLeuLysAAsPLeuG1u 560
Db 1621 GCGTACTTCACTAAGTGAAGAGTGTACCAAGAGAGAAACATATCTGAAGATCTGAA 1680
QY 561 Val11eThrSerTrpPheG1nSerThrVal1SerLysG1uAAsPAlaMeProG1uAlaLeu 580
Db 1681 GTTATCACTTCGTGGTTTCAAGACAGATGACAAAGAGAGAGAGCCGATGCGGAAGCACTG 1740
QY 581 LysSerLeu11ePheProA5nPheG1uProLeuH1eLysPheH1eThraA5nPheLeuLys 600
Db 1741 AAAAGTCTATATTCGGAATTTTGAACCTTTCACAAATTTCACTAATTTTCTCAAG 1800
QY 601 Glu11eG1nArqLeuAlaLeuTrpG1uG1yArqSerA5nAlaG1n11eA5nAAsP1y 620

Db	1801	GAAGATTGAGCAACAGACTTGGCCCTGTGGGAGAAAGCCGCGCTCAAAATGCCAAATCAGAGATTAC	1860
Oy	621	GLNATGIIEGLYABPVALMETLEULYASANIIEGLINGIMETLYENHLEUAAIAHVS	640
Db	1861	CAAGAAATCGGGGATGTGATCGTGAAGAAATTTCAAGGGCATTGAAGCACTGGCGGCTCAC	1920
Oy	641	LEUTRPLYBHSERGLUALALEUGLUALALEUGLUAENGLYILEYSSERSETRATG	660
Db	1921	CTGTGGAAAGCACAGGAGGCGCTTGGAGGCGCTTGGAGAAATGGAAATMACAGCTCCCGGGGG	1980
Oy	661	LEUGLUAANPHECYATGAPRPHEGJULEUGLUNLVVALCYATYLEUPROLEUANTHR	680
Db	1961	CTGGAGAACTTCTGAGAGACCTTTAGCTGCGAAGAGTGTATTACTTACCGCTCAACCC	2040
Oy	681	PHLEULEUAIRGPROLEUHIATRGHEUMETHIATRYLYSGINVALLEUGLUAIRGUCYS	700
Db	2041	TTCCCTCTGGCGGCGCACTGACCGGCTATGCACTACAGAGAGTCTTGAGAGGCTGAC	2100
Oy	701	LYENHBNHSPROPROSERHIALAABRPHEATGAPRCYALGALIALEUAAIGULILE	720
Db	2101	AAACCCACCCGCGCGAGCCACGCGCCACATTCAGGGACTCCGAGCGCTTTGGCAGAGATC	2160
Oy	721	THRGJUMETVALAAGLPLEUHIAGLYTHRMECTILEYMETGUAANPHEGLIPLYAEU	740
Db	2161	ACGGAGATGGTGGCAGACTCCAGCGTACGATGATCAAGATGAGAAATTTCCAGAAAGCTG	2220
Oy	741	HISGLULEULYLYASAPLEULIEGLYILEASPALEUVALVALPROGLYARGLUPHE	760
Db	2221	CACGAACTCAAGAAAGATTGATTGGATTGACATCTTGTGTGTTCCGGGAAGGAGATTC	2280
Oy	761	ILEATRGLEUGLYSERLEUSERLYALEUSERGLYLYSGYLEUGLINGINATGMECPHEPHE	780
Db	2281	ATCCGCTCTGGGCGAGCTCAGACAGCTCTCGGGGAAAGGGGCTCCAGACGCGATGTTCTTC	2340
Oy	781	LEUPHEAANAPVALLEULEUTYCTRHSERARGLYLEUTRPAASERASNGLPHELYS	800
Db	2341	CTGTTCMAACGAGCTCTCTATACACGAGCGGGGGCTGAGCGGCTCCAAATCAGTTTAA	2400
Oy	801	VALHISGLYGLINLEUPPROLEUTYCTGLYMETTHRILEGLUGLUSERGLUAAPGLUTRPSLY	820
Db	2401	GTCCACGGGGCAGCTCCCGCTCTATGCGATTGACATTGAGAGAGCGAAGACAGATGGGGG	2460
Oy	821	VALPROHIVCYALEUTHIRLEUAIRGLYGLINATRGINSERILELEVALIAIASERSE	840
Db	2461	GTGCCCCCACTGCTGACCCCTCCGGGGCCAGCGGCAGTCCATCATCTGTGGCCGCAATGTT	2520
Oy	841	ARGSERGLUMETGLULYETRYVALGLUAAPRIIEGLIMETAAILEAPLEUAAIGLUNLYS	860
Db	2521	CGGTCCGAGATGGAGAAATGGGTGTGAAGACATCCAGATGGCCATTGACTGGCGGAGAAAG	2580
Oy	861	SERSESERPROALAPROGLUPHEULEUAIASERSEPROPOAPPAANLYSESPROAPR	880
Db	2581	AGCAGCAGCCCCCGCCCTGAGTTCCTTGGCCAGCAGCCCCCTGACAAACAGTCCCTGAT	2640
Oy	881	GLUALATHRALAIAAABPGLINGLUSERGLUAAPRPLEUSERALIASERATGTHSERLEU	900
Db	2641	GAAGCAGCCGCGGCTGACACAGAGATCAGAGATGACCTGAGCGGCTCCGCAATCCGCTG	2700
Oy	901	GLIUAIRGINALAPROHIAATRGLYASNTHRMETVALHISVALCYETRPHIASATGANTHR	920
Db	2701	GAGCGCCACGCGCGCCGACCGCGGCAACCAATGGCGACGCTGCTGGCACCGCAACCC	2760
Oy	921	SERVALSERMETVALABPHESESERILEALVALGIUAANGLILEUSERGLYASNLEULEU	940
Db	2761	AGCGTCTTCATGTGTGACATTCAGCATCCGCAATGGAGAAATCAAGTGTCTTGGAAACCTGCTG	2820
Oy	941	ARGLYPHELYEASNSERASNGLYTRPSINLYLEUTRVALVALPHETHRASNPHCYGS	960
Db	2821	AGGAAATTCAAAAACAGCAACGGGGGCGCAAGAGCTGGGGTGGTTCACAAACTTCTGC	2880
Oy	961	LEUPHERPEYTRYBSESEHIEGLINAPRANHSPROLEUALASERLEUPROLEULEUGLY	980
Db	2881	CTGTTCTTCTCAAAATTCACACAGGACATATCCCTTGGCAGGCTGCGCTCTGTGCGG	2940

FEATURES	source
gene	1. .4775 /gene="DKFZp469C092" 100. .3237 /gene="DKFZp469C092" /codon_start=1 /product="hypothetical protein" /protein_id="CAH91292.1" /db_xref="GI:55729111" /translation="MGRIQRPPTGSRILGAPENSGISTLERGQKPPPTPSGLVSIKI OMIDPDAFEVPOPARAGKVLADVAVCNHLAVGSDYGLSCPDKKTITWLDLKLPLV KOIRBPYVYKVFVYKFPFDDLOLOBELTTRYLFAVOVKDLOAGRLTCDTSALLII SHVQSGIDPDEALDBREHLAKNKYITFQDADLEKIVYFHHNIGQTPASDROLLEI ARLEBMYGLRHPKDEGRTKINDAVANTGILVFQGTKNAFWMAVYKLSFGKFF LTKIRPANSAVQDTLEFLMASNDPCKSPFKICVEHNAFRLFEPPKPKPVLFSFGS
CDs	
LOCUS	CR859100 4775 bp mRNA linear HTC 12-NOV-2004
DEFINITION	Pongo pygmaeus mRNA, cDNA DKFZp469C092 (from clone DKFZp469C092).
ACCESSION	CR859100
VERSION	CR859100.1 GI:55729110
KEYWORDS	HTC.
SOURCE	Pongo pygmaeus (orangutan)
ORGANISM	Pongo pygmaeus
REFERENCE	Bukaryote, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi; Mammalia, Eutheria, Euarchontoglires, Primates, Catarrhini, Hominiidae; Pongo. 1 (bases 1 to 4775) Bloecher, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C., Oeinger, A., Fobö, G., Han, M. and Wiemann, S. The German cDNA Consortium Direct Submission Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
CONSRIM	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Sequenced by GBR (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp469C092) is available at the RZPD Deutsches Reisourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/c/ cgi?cloneID=DKFZp469C092 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
JOURNAL	location/Qualifiers
COMMENT	1. .4775 /organism="Pongo pygmaeus" /mol_type="mRNA" /db_xref="taxon:9600" /clone="DKFZp469C092" /tissue_type="kidney" /clone_id="469 (synonym: pkid1). Vector pSPORT1_sfi; host DH10B; sites SfiIA + SfiIB" /dev_stage="adult" /note="FERN, RhocEF, and plectstrin domain protein 1 (Homo sapiens)"

SSPFRSGRTOKOVLVDYVEGGHKKVKYOFERKSKITHSIRSLASOPTELYSEVLBOOS
ALITFRFGASBPGGOSCOQKEPRVSPGSEPSHPSPRBRPAKNKADGASAPTE
EEBVKRTQOSKRPORPORSGLJUSPHLSHSVNSGCVAPRANLSPDSTMO
ASPLISPLNDQACPTDDEDEBGRKRPPTDKAFIAKEVSTERTYTKUDEVTSMF
QSAVSKEDAMPEAKSLIFNFEPLHKFHTNLEKEIQRLALWGRSNAQIRDYQIG
DVMILKNIQGMKGLAVHLMKSEBALALENGIKSRRLBNFORDELOKVCYLPNTFL
LRPLHBMHAYKQVLERLCKHPPSHADFRCDRALAAGITENVAOLHGTIMEMENPOKL
HELKDLIGDNLVYVPGREFIRLSGLSLSGKLOOMFLFNDVLLYTSRGLPNSQ
FKVHGOLPLYMTVTKESEBDEWGVPHCLTLRBOROSITVAASSTRSEMEKMWIDOMAD
LAEKNSSLAPFLASSPPDNKSPDEMTADDESDDLASSTSLERAPRNGRTMAY
CMKRNYSVSWDFVAAVENQJLGNLRFKXNSGMOKLWVFNFPCLPFYXSHDNBP
LASLPLLGYSLTITPTESENIHKDYVFKLHFKSHVYFRASEEYTFERMEVIRSATSS
ASRVHVSHKESLYV"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	4775
Score:	5359.00	Matches:	1025
Percent Similarity:	98.8%	Conservative:	7
Best Local Similarity:	98.1%	Mismatches:	13
Query Match:	98.1%	Indels:	0
DB:	6	Gaps:	0

US-09-555-342B-2 (1-1045) x CR859100 (1-4775)

Qy 1 MetGlyIuIleGluGlnArgProThrProGlySerArgLeuGluYAlaProGluAAsnSer 20
Db 100 ATGGAGAAATRAGACAGAGGCCAAGCCCAAGATCAAGACTGGGGGCCCAAGAAATTCG 159
Qy 21 GlyIleSerThrLeuGluArgGlyGluIleArgProProThrProSerGlyIleLeuVal 40
Db 160 GGGATCAGTACCTTGAAACGTGACAGAACGCCGCCCAACCTTCAGAAATCTCGTG 219
Qy 41 SerIleValIleGlnMetLeuAspAspThrGlnGluIlePheGluValProGluArgAla 60
Db 220 TCCATCAAAATCCAGATGCTGGATGACACCCAGAGGCAATTGAAATTCCAAAGAGCT 279
Qy 61 ProGlyIleValIleLeuAspAlaValCyAsnHisIleLeuAsnLeuValGluIleAspTyr 80
Db 280 CCGTGGAGAGTGTGCTGCTGATGACAGTTTGCAACCACTCAACCTCGTGGAGGTGACTAT 339
Qy 81 PheGlyLeuGluPheProAspHisIleValIleThrValIlePheLeuAspLeuIleArgPro 100
Db 340 TTTGGCTCGAGTGTCTCGATCACAAAAGATCAACGCTGTGGTGGATCTCTTAAATCCC 399
Qy 101 IleValIleGlnIleArgArgProIleHisIleValIleValIlePheValIlePhePhe 120
Db 400 CTTGTGAAACGATCAGAGGSCAAAGCATGTGTGTTAAGTTGTGTGTAATTCCTT 459
Qy 121 ProProAspHisThrGlnLeuGlnGluIleuThrArgTyrLeuPheAlaLeuGlnVal 140
Db 460 CCGGCTGACCAACAACTCCAAAGAAAGATCAACAAGTACCTGTTGCGCTGCAAGTG 519
Qy 141 LysGlnAspLeuAlaGlnGlyArgLeuThrCyAsnAspThrSerAlaAlaLeuIle 160
Db 520 AAGCAGAGCATTTGGCTCAAGGCGAGTTGACGTGTAACGACACAGCGAGCTCTTGATT 579
Qy 161 SerIleIleValIleGlnSerGluIleGlyAspPheArgGluAlaLeuAspArgGluHisIleu 180
Db 580 TCACACATTTGCAATCGAATGGGATTTTGAATGAAGCTTGGACAGAGACACTTA 639
Qy 181 AlaIleAsnIleGlyIleProGlnGlnAspAlaLeuGluAspIleValIleGluPheHis 200
Db 640 GCAAAAATAAATCATCTCAACCAAGACCACTAAGAGCAAAATGTGTGAATTTTCAC 699
Qy 201 HisAsnHisIleGlyIleThrProAlaGlnSerAspPheGlnLeuGluIleAlaArg 220
Db 700 CATTAACCATTTGCAAAACACAGCAAGATCAATTTCCAGCTCTTGAAGATTGCCGT 759
Qy 221 ArgLeuGlnMetTyrGlyIleArgLeuHisProAlaIleAspArgGluGlyIleThrIle 240
Db 760 CGGCTAGAGATGTATGATCCGCTGACCCGGCCAAAGACAGAGAAAGGACAGAAATC 819

Qy 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrIleIleAsnAla 260
Db 820 AATGTGCCCTTGCCAAACAGGAAATTCATGTTTCAAGGTTTCACTAAGATCAATGCC 879
Qy 261 PheAsnThrAlaIleValIleArgIleLeuSerPheIleArgIleValArgPheLeuIleIleLeu 280
Db 880 TTCAACTGGGCCCAAGGCGCGAAGCTGAGCTTCAAGAGAAAGCCCTTTCATCAAGCTC 939
Qy 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAlaSerArgAsp 300
Db 940 CGGCACATCCCAATAGTGGTACAGAGATACCTTGAATTCCTGATGGCCAGTCGGGAT 999
Qy 301 PheCysIleSerPheThrIleIleCysValGluHisIleAlaPhePheArgLeuPheGlu 320
Db 1000 TTCTGCAAGTCTCTTGAAATCTGTGTTGAAACATCATGCTTCTTAACTTTTGAA 1059
Qy 321 GluProIleArgProIleProIleProValLeuPheSerArgIleSerSerPheArgPheSer 340
Db 1060 GAGCCCAAAACCAAGGCCCAAGCCCGTCTCTTAAAGCCGGGCTCATATTCGGTTCAGT 1119
Qy 341 GlyArgThrGlnIleGlnValIleuAspTyrValIleGluGlyHisIleIleValIleGln 360
Db 1120 GGTGGACTGAGAAAGCAGTTCTCGACTATGTTAAGAGAGAGACATTAAGAGGTGAC 1179
Qy 361 PheGluArgIleHisIleSerIleIleHisIleArgSerLeuAlaSerGlnProThrGlu 380
Db 1180 TTTGAAAGAGACACAGCAAGATTCATTCATCCGAGCCTTGCTTCAAGCCTTACAGAA 1239
Qy 381 LeuAsnSerGluValIleuGlnIleSerGlnIleSerThrSerLeuThrPheGlyGlnGly 400
Db 1240 CTGTATTCGAAAGTACTGGAGAGTCTCAGACAGCCCAAGCCTTCAATTTGGAGAAAGT 1299
Qy 401 AlaGluSerProGlyGlnIleSerCyAsnArgGlyIleGlnProIleValIleSerAlaGly 420
Db 1300 GCCAATCTTCAAGGGGCCAGAGCTGCCACAGAAAGAAACCAAGATTTTCCCGGGG 1359
Qy 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGluAsnIleGlnAla 440
Db 1360 GAGCCGGGGTGGACCCGAGCCCTGTGCGAGAGAAAGCCCGGGGTAAACAAGCAGCG 1419
Qy 441 AspGlyAlaIleSerAlaProThrGluGluGluGluValIleValIleAspArgThrGln 460
Db 1420 GACGAGACCCCTCGGGCCCAAGAGAGAGAGAGGTGTTAAGATAGGACCAAG 1479
Qy 461 GlnSerIleProGlnProProGlnProSerThrGlySerLeuThrArgIleSerProHisIleu 480
Db 1480 CAGATTAACCTCAGCCCGGACGACCAAGCAGAGCTCCTGACCGGACGCTCACCTT 1539
Qy 481 SerGluLeuSerValaAsnSerGlnGlyIleAlaIleProAlaAsnValIleLeuSerPro 500
Db 1540 TCTGAGCTGTCTGAACTCGAGAGGGGAGTGGCCCTGCCAAGCTGACTGTCTCC 1599
Qy 501 AsnLeuSerProAspThrIleGlnAlaSerProLeuIleSerProLeuLeuAsnAspGln 520
Db 1600 AACCTGAGCCCAACACAGAGGCTCTCCCTGATACGCCGCTGCTGAAGACACAG 1659
Qy 521 AlaCysProArgThrAspAspGluAspGluGlyIleArgIleIleArgPheProThrAspLys 540
Db 1660 GCTGCCCCCGGACGACGATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATAA 1719
Qy 541 AlaIlePheIleAlaIleGluValIleSerThrThrGluArgThrTyrLeuIleAspLeuGlu 560
Db 1720 GCGTACTTCACTAAGTAAGAAAGTGTCAACCAAGAGGAACATATTCGAAGATCTGAA 1779
Qy 561 ValIleThrSerThrPheGlnIleSerThrValIleSerIleGluAspAlaMetProGluAlaLeu 580
Db 1780 GTTATCACTTCGGTTCAGAGCGCAGTGAAGAAAGAGACGCGCAAGCAAGCACTG 1839
Qy 581 LysSerLeuIlePheProAspPheGluProLeuHisIleIlePheHisIleThrAsnIleLeu 600
Db 1840 AAAGTCTCATATTCGCCAATTTTGAACCTTTTGCAAAATTTTCACTACCAATTTTCTAAG 1899
Qy 601 GluIleGlnIleArgLeuAlaLeuTyrGluGlyArgSerAsnAlaGlnIleArgAspTyr 620

```

Db      1900 GAAATTGACAGCACTTGCCTGTGGAGGCGCCTCAAAATGCCAATACGAGATTAC 1959
Qy      621 Glnatgllleglyaapvalmetleulyabnllleglmglymetlyhbleuallahis 640
Db      1960 CAAGAATCGCGGAGTGTATGCTGAAGAACATTCCAGGGGATGAAACACCTGCGGTTAC 2019
Qy      641 Leutrpbyhshserglualaleuglualaleugluabnglylaleyserserargx 660
Db      2020 CTGTGAAAGACAGCGAGGCTTGGAGGCGCTGGAAGATGGATCAAGAGCTCCGCGCG 2079
Qy      661 Leugluabnphocyargapbphelileuglnlyvalcyetyrleuproleuabnthr 680
Db      2080 CTGGAGAACTTCTGAGAGACTTTAGCTGCGAAGAGTGTGTACTACCGCTCAACACC 2139
Qy      681 Pheuleuargproleuhihargleuethistyrlyrvglnvalleugluargleucys 700
Db      2140 TTCCCTCTGCGGCGCCTGCACCGGCGCATGCACTCAAGCAAGGTCTTGAGCGGCTGTGC 2199
Qy      701 LysHshhshppropserhshalabpPheargapcyargalalaleuallagluile 720
Db      2200 AAAACACACCGCGCAGCGACGCGCATTCAGGGAGCTGCGAGCGCTTGCGAGGAGATC 2259
Qy      721 Thrglumetvalalaglnleuhihgllythrmctileysmetgluabnphelinybleu 740
Db      2260 ACGGAGATGATGGGCACTCCACGCGTACGATCAAGATGAGAGATTTCCAGAGAGCTG 2319
Qy      741 Hlsgleuleulyayapbpleuileglylileapabnleuvalvalproglyargglubhe 760
Db      2320 CACGAACTCAAGAAAGATTGTGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 2379
Qy      761 IlleargleuglyserleuSerleuSerlyleuSerglyyleuglnlarymerphehe 780
Db      2380 ATCCGTCTGGGAGGCTCTGAGCAAGCTCTGGGGAAGGGGCTCCAGACGCGCATGTTCTC 2439
Qy      781 LeupheaenapvalleuleuTytrhserargglyleuThnlaserabnglnphelys 800
Db      2440 CTGTTCACAGACGCTCTCTGTACACAGCCGCGGCTGACGCGCTCCATCAGTTTAA 2499
Qy      801 ValhshglglnleuproleuTyrglymetthrllleglglusergluabpglutrpily 820
Db      2500 GTCCACGCGCAGCTCCGCTCTCAACGCGCATGACGATTAAGAGAGCGAAGACAGCGGCGG 2559
Qy      821 ValprohshcybleuThleuargglylgnarghlnserlilevalalalaser 840
Db      2560 GTGCGCCACCTGCTACCTCCGGGCGCAGCGCGCATCATCATGTTGCGCGCGCATCT 2619
Qy      841 Argserglumetglulyetrpvalgluabpilleglnmetalaleapleuallagluys 860
Db      2630 CGGTCTGAGATGAGAAAGTGGGTTAGAGCATCCAGATGCGCATTCGCTGCGAGAGAG 2679
Qy      861 SerSerSerProalaprogluPheleuallaserSerProProapbapnlyserProap 880
Db      2680 AACAGCACCTCTGCGCTCTGAGATTCTGTGCGACAGACCCCTGACACACAGTCCCTGAT 2739
Qy      881 Gluualatrralalaaapglglusergluabpabnlyaserargthrsereu 900
Db      2740 GAAGCACCGCGCTGACAGAGATCAAGATGACCTGAGGCGCTCGCGCACATCGCTG 2799
Qy      901 Gluargglinalaprohshargglyabnthrmctvalhshvalcyetyrphshargabnthr 920
Db      2800 GAGCGCCAGCGCCCGCAGCTGCGCAACATGTGTGATGTGTGTGCGACCGCAACACC 2859
Qy      921 SerValSerMetValapbpheserllealavalgluabnglnleuserglyabnleu 940
Db      2860 AGCGTCTCATGTGTGATTCAGCTGACGTGAGTGAAGATCAAGTTGTCTGGGAACCTGCTG 2919
Qy      941 ArglyvshelyvabnserabnglytrpnllybleuTytrpvalvalPheThPabnphocy 960
Db      2920 AGGAAATTCAAAGACGACGCGTGGCAAGAGCTGTGGGTGTGTTCACAAATCTTGC 2979
Qy      961 Leuphepetyrlyserhshglinaapbapnshproleuallaserleuproleuugly 980

```

```

Db      2980 CTGTCTCTTAAGTACACACGAGCAATCATCCCTCGCAGCTGCTGTGCGGC 3039
Qy      981 TyserleuThrlleProSerGlusergluabnllleglmglyyapbpyvalPheyleu 1000
Db      3040 TACTGCTCACCATCCCGCATGAGTCCGAGAACATCCCAAGAGTACGTGTCAAGCTG 3099
Qy      1001 HlshpelysSerhshvallyrtyrPhearglallagusergluTytrhPhegluargtrp 1020
Db      3100 CACTTCAAGTCCACAGTACTTACTTCAGGCGGAAACGAGTACACTTCGAAAGTGG 3159
Qy      1021 MetglualllarySerValaThrserserlaserargprohshvalleuSerhshlys 1040
Db      3160 ATGGAATGATCCGAGCGCGCACGCTGCTGCGAGTCCAGTCCAGTGTGAGCCACAAA 3219
Qy      1041 GluserleuValtyr 1045
Db      3220 GAGTCTCTCGTGTAT 3234

RESULT 3
AY406168
LOCUS      AY406168      3138 bp      DNA      linear      GSS      15-DEC-2003
DEFINITION Pan troglodytes PARP1 gene, VIRUTAL TRANSCRIPT, partial sequence,
ACCESSION AY406168
VERSION   AY406168.1 GI:39762142
KEYWORDS  GSS.
SOURCE    Pan troglodytes (chimpanzee)
ORGANISM  Pan troglodytes
REFERENCE 1 (bases 1 to 3138)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
          Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
          Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
          Adams,M.D. and Cargill,M.
          Inferring nonneutral evolution from human-chimp-mouse orthologous
          gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED    14671302
REFERENCE 2 (bases 1 to 3138)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
          Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
          Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
          Adams,M.D. and Cargill,M.
          Direct Submission
          Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
          Rockville, MD 20850, USA
          This sequence was made by sequencing genomic exons and ordering
          them based on alignment.
FEATURES
     source
         1..3138
             location/Qualifiers
                 /organism="Pan troglodytes"
                 /mol_type="genomic DNA"
                 /db_xref="taxon:9598"
                 <1..3138
                 /gene="PARP1"
                 /locus_tag="HCM2467"
ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      3138
Score:          5116.00      Matches:      989
Percent Similarity: 94.8%      Conservative: 2
Best Local Similarity: 94.6%      Mismatches: 54
Query Match:    93.6%      Indels:      0
DB:             14      Gaps:      0
US-09-555-342B-2 (1-1045) x AY406168 (1-3138)
Qy      1 MetglualllleglgnargProThrProglySerarglueugllyalaprogluabnser 20
Db      1 ATGGAGAAATPAGACAGAGCGCACCCAGAGATCGGACTGGGGGCGCCCGGAAATTCG 60

```

Qy 21 GlyIleSerThrLeuGluArgGlyGlnIleYsPProFProThrProSerGlyIleLeuVal 40
Db 61 GGGAATCGATGCTTGGAAAGCGTGAACAGAGCCGCCCAACACCTTCAGAAATCTGCTG 120
Qy 41 SerIleYsIleGlnMetLeuAspSerThrGlnGluAlaPheGluValProGlnArgAla 60
Db 121 TCCATCAAAATCCAGATGCTGTGATGACACCCAGAGGCGCATTTGAAGTTCCANNNNNNNN 180
Qy 61 ProGlyYsValLeuLeuAspAlaValCysAsnHisIleuAsnLeuValGluGlyAspTyr 80
Db 181 NNN 240
Qy 81 PheGlyLeuGluPheProAspHisIleYsIleThrValTyrLeuAspLeuLeuYsPro 100
Db 241 NNN 300
Qy 101 IleValIleYsGlnIleArgArgProYsHisIleValIleYsPheValValIleYsPhe 120
Db 301 ATTGTGAACAGATTAGAGAGGCCAAAGCATGTTGTGTTAAGTTGTGTGAAATTTCTTT 360
Qy 121 ProProAspHisSerThrGlnIleuGlnGluIleuThrArgTyrIleAspPheAlaLeuGluVal 140
Db 361 CCGCTGACCAACACAACTCCAAAGAAAGACTCAAGAGTACTGTTCCGCTCAGAGTG 420
Qy 141 LysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAlaAlaLeuLeuIle 160
Db 421 AAGCAGAGCTTGGCTCAAGCAGGTTGACGTGTAATGACACACAGCAGGCTCTTGATT 480
Qy 161 SerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgGlnHisIleu 180
Db 481 TCACACATTTGCAATCTGAGATTGGGGATTGATGTAAGCCTTGAGCAGAGACACTTGA 540
Qy 181 AlaIleYsAsnIleYsTyrIleProGlnGlnAspAlaLeuGlnAspYsIleValGluPheHis 200
Db 541 GCAAAATTAATATCATRACCTCAGCAAGACCCACTTAAGACAAATCGTGGAAATTTTAC 600
Qy 201 HisAsnHisIleGlyGlnThrProAlaGlnSerAspPheGlnIleuLeuGluIleAlaArg 220
Db 601 CATACCACTTGGCAAAACACAGCAAGAAATCATTTCCAGCTCTTAAGATTGGCCGT 660
Qy 221 ArgLeuGlnMetTyrGlyIleArgLeuHisProAlaIleYsAspArgGluGlyThrLysIle 240
Db 661 CCGGTAGAGATGTATGGAAATCCGGTTCACCCGCGCAAGACAGAGGACCAAAATC 720
Qy 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAla 260
Db 721 AATCTGCGCTTGGCAACAGGAAATTTCTAGTGTTCAGGGTTTCACTTAAGATCATGCC 780
Qy 261 PheAsnTyrAlaYsValArgYsLeuSerPheYsArgYsArgPheLeuIleYsLeu 280
Db 781 TTCAACTGGGCGCAAGGTCGAGACTGAGCTTCAAGAGGAAGCGCTTTCATCATAGCTC 840
Qy 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrIleuGluPheLeuMetAlaSerArgAsp 300
Db 841 CCGGCATATGCAATAGTGGCTACCAAGATACCTTGAATTTCCGATGGCCAGTCCGAGAT 900
Qy 301 PheCysLeuSerPheTyrPlyIleCysValGlnHisIleAlaPheArgLeuPheGlu 320
Db 901 TTCTGCAAGTCTTCTGAAATATGTGTGTGAACATCATGCTTCTTTAAGCTTTTGA 960
Qy 321 GluProIleYsProIleYsProValLeuPheSerArgIleYsSerPheArgPheSer 340
Db 961 GAGCCCAACCAAGCCCAAGCCCGTCTTTAGCCGGGGGTATCATTTCCGTTCACT 1020
Qy 341 GlyArgThrGlnIleYsGlnValLeuAspTyrValIleGluGlyGlyHisIleYsValGln 360
Db 1021 GGTCGAGCTCAGAGAGGTTCTCGACTATGTAAAGAGAGGACATTAAGAGGTCCAG 1080
Qy 361 PheGluArgYsHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
Db 1081 TTTCAAAGGAAGCAGCAGATTCATTCTATCCGAGCCTTGCTTCAACGCTTACAGAA 1140

Qy 381 LeuAsnSerGluValLeuGlnIleuSerGlnIleSerThrSerLeuThrPheGlyGly 400
Db 1141 CTGAATTCGGAAAGTGTGAGAGAGTCTACAGAGGCCCAAGCTTATCATTTGAGAAAGT 1200
Qy 401 AlaGlnSerProGlyGlyGlnIleuSerCysArgArgGlyYsLeuProYsValSerAlaGly 420
Db 1201 GCCGAATCTCCAGGGGGCCAGAGCTGCGGCAAGAAAGAAACGGAAGTTTCCGCGGG 1260
Qy 421 GluProGlySerHisPProSerProAlaProArgArgSerProAlaGlyAsnYsGlnAla 440
Db 1261 GAGCCGGGGTTCGACCCGAGCCCTGCGCCGAGAGAAAGCCCGCGGTAAACAGACAGCG 1320
Qy 441 AspGlyAlaAlaSerAlaProThrGlnGluGluGluValIleYsAspArgThrGln 460
Db 1321 GACGAGCCGCTCGGGGCCACGAGAGAGAGAGAGAGTCCGTTAAGATAGGACCCAG 1380
Qy 461 GlnSerIleYsProGlnProProGlnProSerThrGlySerIleuThrGlySerProHisIleu 480
Db 1381 CAGAGTAAACCTAGCCCGCCGAGCCCAAGCAGAGCTCCCTGACTGGCAGTCTCACCTT 1440
Qy 481 SerGluSerValAlaAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro 500
Db 1441 TCCGAGCTGTCTGAACTMGACAGGGGGAGTGGCCCTTCGCAAGCTTCTCTCC 1500
Qy 501 AsnLeuSerProAspThrIleYsGlnAlaSerProLeuIleSerProIleuAsnAspGln 520
Db 1501 AACCTGAGCCCGACACCAAGCAGGCTCTCCCTGATCAGCCGCTGTAATGACAG 1560
Qy 521 AlaCysProArgThrAspArgIleuAspGluGlyArgArgYsArgPheProThrAspLys 540
Db 1561 GCGTCCCCCGAGCGAGCATGAGATGAGAGGCGCGAGGAAGATTTCCAACTGAATAA 1620
Qy 541 AlaTyrPheIleAlaYsGluValSerThrThrGlnArgThrTyrLeuYsAspLeuGlu 560
Db 1621 GCGTACTTCACTAGCTAAAGAGNNNTCCNNNACCGAGGAAACATATCGAAGATCTTGA 1680
Qy 561 ValIleThrSerTyrPheGlnSerThrValSerIleGluAspAlaMetProGluAlaLeu 580
Db 1681 GTTATCACTTCGTGGTTTCAGACACAGTAGCAAAAGAGACCCATGCGCGAAGCACTG 1740
Qy 581 LysSerIleuIlePheProAspPheGluProLeuHisIleYsPheHisIleThrAsnIleYs 600
Db 1741 AAAAGTCAATATCCGAATTTGAACCTTTCGCAAAATTTCAATATTAATTTCTCAAG 1800
Qy 601 GluIleGlnIleArgLeuAlaLeuTyrGluGlyArgSerAsnAlaGlnIleArgAspTyr 620
Db 1801 GAATATGAGCAAGACTGCTGCCCTGGGAAGGCGCTCAATATGCCAAATCAAGATTAAC 1860
Qy 621 GlnArgTleGlyAspValMetLeuYsAsnIleGlnGlyMetYsIleuAlaAlaHis 640
Db 1861 CAAGAATCGCGCATGTCACTGAAGAACATTCAGGGCATGAAGCACCTGGCGGCTCAC 1920
Qy 641 LeuThrYsHisSerGluAlaLeuGluAlaLeuGluAsnGlyIleYsSerSerArgArg 660
Db 1921 CTGTGAAGACACAGAGGCTCGAGGCGCTGAGAGTGAATGAATCAAGACTCCCGCGG 1980
Qy 661 LeuGlnAsnPheCysArgAspPheGluLeuGlnIleValCysTyrIleuProLeuAsnThr 680
Db 1981 CTGGAAGACTTCTGCAAGACTTGAAGCTGCAAGAGTGTCTTACACCGCTCAACACC 2040
Qy 681 PheLeuLeuArgProLeuHisIleArgLeuMetHisTyrIleGlnValLeuGluArgLeuCys 700
Db 2041 TTCTCTCTGGCGCAGTGCACCGGCTCATATGCANNACAAAGAGGCTCGAGCGGNNGTGC 2100
Qy 701 LysHisIleProProSerHisIleAspPheArgAspCysArgAlaAlaLeuAlaGluIle 720
Db 2101 AAACACACCCCGCGGANN 2160
Qy 721 ThrGlnMetValAlaGlnIleuHisGlyThrMetIleYsMetGluAsnPheGlnIleYsLeu 740
Db 2161 ACGAGATGGTGGCAGACGCTCCAGCGGTACGATATCAAGATGAGAAATTTCCAGAACTG 2220
Qy 741 HisGluLeuYsIleAspLeuIleGlyIleAspAsnLeuValIleProGlyArgGluPhe 760

```

Db      2221 CACGAACCTCAAGAAAGATTGATGTCATGACATCTTGCGTCCGGGAAAGGAGTTC 2280
Qy      761  ILeArgLeuGlySerLeuSerLeuSerGlyLyLeuGlnGlnArgMetPhePhe 780
Db      2281 ATCCGCTCTGGGACAGCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGCACGCGATGTTCTTC 2340
Qy      781  LeuPheAaAPValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsnGlnPheLyS 800
Db      2341 CTGTTCAACGACGCTCTCTATACACGAGCCGGGGGCTGACGGCTCCAAATCAGTTTAA 2400
Qy      801  ValHLeGlyGlnLeuProLeuTyrGlyMetThrIleGlnGlnSerGlyAaAPGlyL 820
Db      2401 GTCCACGGGACAGCTCCGCTCTATGCGATGACATGAGACAGCAAGGAAAGAGGGGG 2460
Qy      821  ValProHisGlyLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAsnSer 840
Db      2461 GTGCCCACTGCTGACCTCCGGGGCCAGCGCGAGTCATCATCTGCGCCGCGAGTCT 2520
Qy      841  ArgSerGluMetGluValTyrValGluAaPrlIleGlnMetAlaIleAsnLeuAgl 860
Db      2521 CGGTCCGAGATGAGAAATGGGTGAGAGACATCCAGATGGCCATTGACCTGGCGAGAG 2580
Qy      861  SerSerSerProAlaProGluPheLeuAlaSerSerProProAsnLySerProAsp 880
Db      2581 AGCAGCAGCCCCCGCCCTGAGTCTCTGGCCAGCAGCCCCCTGACAAACAGTCCCTGAT 2640
Qy      881  GluAlaThrAlaAlaAsnArgGlnGlySerGluAaPheLeuSerAlaSerArgThr 900
Db      2641 GAACGACCGCGGGCTGACCAAGAGTCAAGAGATGACCTGACGCGCTCGCGCACGTCG 2700
Qy      901  GluArgGlnAlaProHisArgGlyAaSerMetValIleValCysThrHisArgAsn 920
Db      2701 GAGGCGCAGGCGCCGCGACCGGCAACAAATGGTGACGTGCTGGCCACCGCAACCC 2760
Qy      921  SerValSerMetValaAsnPheSerIleAlaValGluAaGlnLeuSerGlyAa 940
Db      2761 AGCGTCTCAGTGGTGAATTCAGCATCCGACGTGGAGATCAGTGTCTCGGAAAC 2820
Qy      941  ArgGlyPheLeuAaAsnSerAsnGlyTyrGlnLyLeuTyrValValPheThrAsn 960
Db      2821 AGGAAATTCAAAAACAGCAACGGGTGGCGAAGCTGTGGTGTGTTCAAAACTTCTGC 2880
Qy      961  LeuPhePheTyrLySerHisGlnAsnAsnHisPheLeuAlaSerLeuProLeu 980
Db      2881 CTGTTCTTCAAGTCAACGACACACGAGCAATATATCCCTTGGCAAGCTCTGCTG 2940
Qy      981  TyrSerLeuThrIleProSerGlySerGlyLeuAsnIleGlnLyAsnTyrVal 1000
Db      2941 TACTCGCTCACATCCCTCCGAGTCCGAAACATCCGAAAGACTACGTGTTCAAGCTG 3000
Qy      1001 HisPheLySerHisValTyrTyrPheArgAlaGlnSerGlyTyrThrPheGln 1020
Db      3001 CACTTCAAGTCCCAAGCTTACTTACTTCAAGGGCGGAAACGAGTACAGTTGAAAG 3060
Qy      1021 MetGluValIleArgSerAlaThrSerSerAlaSerArgProHisValLeuSer 1040
Db      3061 ATGAAAGATCCGAGTGCACACGACGACGACGACGACGACGACGACGACGAC 3120
Qy      1041 GluSerLeuValTyr 1045
Db      3121 GAGTCTCTGCTGAT 3135

```

```

RESULT 4
LOCUS      AY406169          3147 bp    DNA          linear    GSS 15-DEC-2003
DEFINITION Mus musculus FARP1 gene, VRUTUAL TRANSCRIPT, partial sequence,
ACCESSION AY406169
VERSION    AY406169.1  GI:39762143
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus

```

```

REFERENCE
AUTHORS    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
            Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE      Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL    Science 302 (5652), 1960-1963 (2003)
PUBMED     14671302
REFERENCE  2 (bases 1 to 3147)
AUTHORS    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
            Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
            Adams,M.D. and Cargill,M.
TITLE      Direct Submission
JOURNAL    Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT    This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES
            source
            location/Qualifiers
            1..3147
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            <1..>3147
            /gene="FARP1"
            /locus_tag="HCM2467"
ORIGIN
Alignment Scores:
Pred. No.:      0          Length:      3147
Score:          5054.50    Matches:      962
Percent Similarity: 94.8%    Conservative: 32
Best Local Similarity: 91.8%    Mismatches: 51
Query Match:    92.5%      Indels:      3
DB:              14          Gaps:      2
US-09-555-342B-2 (1-1045) x AY406169 (1-3147)
Qy      1  MetGlyGlnIleGlnGlnArgProThrProGlySerArgLeuGlyAlaProGluAaAsnSer 20
Db      1  ATGGAGAAATAGAGCAGAAAGCCACCCACATCATGCTCTGGGGGCCGAGAAACTCG 60
Qy      21  GlyIleSerThrLeuGluArgGlyGlnLyAsnProProThrProSerGlyLyLeuVal 40
Db      61  GGATCAGCACCTTGAAACGAGACAGAAACACCGCCACACTTCAGAAAACCTTATG 120
Qy      41  SerIleLyIleGlnMetLeuAaAsnThrArgGlnGlnAlaPheGlnValProGlnArgAla 60
Db      121  ACTGTCAAAATTCAGATGCTGAGACACCCAGAGAGGCTTTGAAGTTCCACACAGACG 180
Qy      61  ProGlyLyValIleLeuAaAsnAlaValCysAsnHisIleAaAsnLeuValGlnGlyAsnTyr 80
Db      121  ACTGTCAAAATTCAGATGCTGAGACACCCAGAGAGGCTTTGAAGTTCCACACAGACG 180
Qy      181  CAGGAGAAAGTTCGTTGTAACGCGGTGCAACCACTCAACTTGTGAGAGGCGACTAC 240
Db      81  PheGlyLeuGlnPheProAsnHisLyAsnIleThrValTyrPheAsnLeuLyAsnPro 100
Qy      241  TTCGGCTGAGATTCCCTGACCAACAGAAAGATCGGTGGTGGATCTCTGAAAGCCC 300
Db      101  IleValLyGlnIleArgArgProLyHisAlaValValLyPheValValLyPhePhe 120
Qy      301  ATTGTGAAGCAGATTCCGAAGACGACGACGACGACGACGACGACGACGACGACGAC 360
Db      121  ProProAsnHisThrGlnLeuGlnGlnGlyLeuThrArgTyrLeuPheAlaLeuGlnVal 140
Qy      361  CTTCCCAACACACACAGCTCCAGAAAGAACTTACAGAGTACCTGTTGCTCTCAGAGTG 420
Db      141  LysGlnAaPheLeuAlaGlnGlyArgLeuThrCysAsnAsnThrSerAlaIleLeuLeu 160
Qy      421  AAGCAGAGCTTGGCTCAAGGCGAGGCTGACGTAAACGACACAGAGCGGCTCTCTGATC 480

```

QY 161 SerHisIleValIleGlnSerGluIleGIYAAspPheAspGluAlaLeuAspArgGluHisLeu 180
DB 481 TCACACATTGACGTCGAGATGGGATTTCCATGAGCTTTGGACAGAGAACCTTA 540
QY 181 AlaIysAsnIleYrTrieProGlnGlnAspAlaLeuGluAspLysIleValIleGluPheHis 200
DB 541 GCAAAAATAATAGTACCTCAACAGGATCGCTGGAGACAGAAATCATGAGATTCCAC 600
QY 201 HisAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuGlnIleAlaArg 220
DB 601 CACGCCACGTTGGACAAACCCCGGACAGTCAATTTCCAGCTTCAGAGGTGGCCGT 660
QY 221 ArgLeuGlnMetYrGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysIle 240
DB 661 CGGCTTAGATGTCGGAATCGGTTACACCTTCGACAGACAGAGAGGACCAAGATC 720
QY 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAla 260
DB 721 AACCTGGCTGTCCCAACACTGGAATTCAGTGTTCAGAGCTTCACGAAAGATCAATGCT 780
QY 261 PheAsnThrAlaLysValArgLysLeuSerPheLysArgLysArgPheLeuIleLysLeu 280
DB 781 TTTAACGGGCGCAAGGTGGGAACTGACTCAAGCGGAAAGCCCTTCTTATCAAGCTC 840
QY 281 ArgProAspAlaAsnSerAlaYrGlnAspThrLeuGluPheLeuMetAlaSerArgAsp 300
DB 841 AGACCTACCGGAATAGCTCATACAGACACCTTGAAGTTCTCGATGGCGGCGGAGAC 900
QY 301 PheCysLysSerPheThrLysIleCysValGlnHisIleAlaPheArgLeuPheGlu 320
DB 901 TTTCGCAAGCTCTTCTGAAATAATGTGTGGACACCATGCTTCTTCAAGCTCTTGAA 960
QY 321 GluProLysProLysProLysProValLeuPheSerArgLysSerSerPheArgPheSer 340
DB 961 GAACCTAAGCCAAAGCCAAAGCCGTCTCTTTCAGCCGAGGGTGCTCTCCGGTTAAGT 1020
QY 341 GlyArgThrGlnLysGlnValLeuAspTyrValLysGlnGlyGlyHisLysLysValGln 360
DB 1021 GGTGCGACTCAGAAAGCAAGTTCTGATTATGTTAAAGCGGCGCCACAAAGAGTCCAG 1080
QY 361 PheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
DB 1081 TTTCGAAAGAAAGCACACCAAGATCCATCCACAGAGCCGTGCTCAAGCCCAAGCA 1140
QY 381 LeuAsnSerGluValIleGlnGlnSerGlnGlnSerThrSerLeuThrPheGlyGlnGly 400
DB 1141 CCGAATTCAGAAAGTGCAGAAACAGTCTCCACAGAGCCCAAGCCTTAGCTTGGAGAGT 1200
QY 401 AlaGluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGly 420
DB 1201 ACCGAGCTCCGGGTGGCCAGAGCTGCCAGCAACGAAAGAAACAAAGCCTGACCTTG 1260
QY 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAla 440
DB 1261 GAGTGGGAGACTCCCAAGAGCCCGCCCTCCCTCAAGAGCCCTCCAGGAGAGCAAGTGA 1320
QY 441 AspGlyValAlaAsnAlaProThrGlnGlnGlnGlnValVal-----LysAspArg 458
DB 1321 GATGGAACCAACCGTGGCCGCTCGAGAGAGAAAGAGAGAGGGGTGCGCAAGAGCGC 1380
QY 459 ThrGlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerPro 478
DB 1381 ATCCGGGCAAGTAACTCAGCCCGCCGACAGCAAGCAAGGCTCCCTGACTGTGAGCCT 1440
QY 479 HisLeuSerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeu 498
DB 1441 CACCTTCAAGAGCTGTCACTCACTCAAGGAGAGAGCCCGCCAGCAAGTGAACCTTG 1500
QY 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuLeuAsn 518
DB 1501 TCTCCCAACCTGAGTCTGACAAACAGAGGCTCTCCCTGATGACAGCCGCTCTGTAAC 1560

QY 519 AspGlnAlaCysProArgThrAspAspGluAspGluGlyValArgArgLysArgPheProThr 538
DB 1561 GACCAAGCGTCCCAAGAGACCGAGACGAGAGAGAGGCGCGCAAGAGATTTCCCAACC 1620
QY 539 AspLysAlaYrPheIleAlaLysGluValSerThrThrGluArgThrYrLeuLysAsp 558
DB 1621 GACAGGCGGTACTACATTTGCTAAGAAAGTCTCCACCAAGAGCGCACGTCAAGAGAC 1680
QY 559 LeuGluValIleThrSerThrPheGlnSerThrValSerLysGluAspAlaMetProGlu 578
DB 1681 CTGAGGTCAATGCTTGTGTGTGTTCAGAGCAAGTCAAGAAAGAGACTCCATGCCGAA 1740
QY 579 AlaLeuLysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAsnPhe 598
DB 1741 GCCTTGAAGAGTCTCAATTTCCAAATTTGAACCTTTGCAAGATTCCTATCAATTTT 1800
QY 599 LeuLysGlnIleGlnGlnArgLeuAlaLeuTyrGlnGlyValArgSerAsnAlaGlnIleArg 618
DB 1801 CTGAAAGAAATTGAGCAACGACTTGCCTGGGAAAGGCGGCTCAATGCCCATATGAGA 1860
QY 619 ---AspTyrGlnArgIleGIYAAspValMetLeuLysAsnIleGlnGlyMetLysHisLeu 637
DB 1861 GGAATTAACCAAGAAATCGGGAGCTCATCTGAAAACATTCAGGGCAATGAGCACCTTG 1920
QY 638 AlaAlaHisLeuTyrLysHisSerGluAlaLeuGluAlaLeuGluAsnGlyIleLysSer 657
DB 1921 GCGGCCACCTGTGGAAAGCAACATGAGGCGCTGAGAGCCCTGGAGACCTCATCAAGGCG 1980
QY 658 SerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuPro 677
DB 1981 TCCGGGGCTGGAAACATTTCTGCCAGACTTGAGAGTGAGAGGTGTTCCTGCGC 2040
QY 678 LeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisTyrLysGlnValIleGln 697
DB 2041 CTCAACACTTCTCTCTGCGCCCGCTGCACCGGCTCATGCACTTAAGCATGCTCGAG 2100
QY 698 ArgLeuCysLysHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeu 717
DB 2101 AGCGTGGCAAGACCAACCAACCAACAGCGGACTTCAGGGAGCTGCAGAGCTGGCTG 2160
QY 718 AlaGluIleThrGluMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPhe 737
DB 2161 GCGAGATTCACAGAGATGGTGGGCAAGCTGCATGATACCTGATCAAGATGAGAGAACCTTC 2220
QY 738 GlnLysLeuHisGlnLeuLysLysAspLeuIleGlyIleAspAsnLeuValProGly 757
DB 2221 CAGAGCTGCATAGACTCAAGAAAGATCTGATGGCATTAACAATCTTGATCCAGAGA 2280
QY 758 ArgGluPheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArg 777
DB 2281 AGGGAATTCATCCGCGTGGCAAGCTTCAGCAAGCTCCGAGAAAGGGGCTTCAGACGCG 2340
QY 778 MetPhePheLeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsn 797
DB 2341 ATGTCTTCTCTGTTCAACGATGTCTTGCTGTATACAGCGGGGAGCTGAGGCAATCTAAT 2400
QY 798 GlnPheLysValHisGlyGlnLeuProLeuTyrGlyMetThrIleGlnGlnGlnGluAsp 817
DB 2401 CAGTTTAAGTCCAGGACAGCTCCCATCTTAAGCATTAAGCATTAAGAGAGTGAAGAG 2460
QY 818 GluTyrGlyValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAla 837
DB 2461 GAGTGGGGTGTGCCCATGCTGACCCGTGCGGGGCAAGGCAATCTATCATCTGTGGCT 2520
QY 838 AlaSerSerArgSerGluMetGlnLysTyrValGluAspIleGlnMetAlaIleAspLeu 857
DB 2521 GCCAGTTCCTCGGTCAAGATGGAATAATGATGAGAGACATCAAGATGGCATTTGACCTG 2580
QY 858 AlaGluLysSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLys 877
DB 2581 GCAAGAAAGCAATGAGACCCACCGAGCTCTTGCCAGAGAGCCCGCTGACAAAG 2640
QY 878 SerProAspGluAlaThrAlaAlaAspGlnGlnSerGluAspAspLeuSerAlaSerArg 897

Db 2641 TCCCGGATGAGCCACAGCGGCTGACCAAGAGTCAAGAGTACGTAGTCCCTCTCGG 2700

Qy 898 ThrSerLeuGluArgGlnAlaProH1sArgGlyAanThrMetValH1sValCysTrpHis 917

Db 2701 ACATGCTGAGAGCGCCAGGCGCCCTCAACCGTGCACACAGATGTGCACGTGTGGCAT 2760

Qy 918 ArgAanThrSerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerGly 937

Db 2761 CGCAGACACAGAGTGTCTCCATGTGTGAGACTTCACATTCCTGCTGGAGAACACAGCTCTCTGGG 2820

Qy 938 AsnLeuLeuArgGlyPhePheAsnSerAanGlyTrpGlnLysLeuTrpValAlaPheThr 957

Db 2821 AACCTGCTGAGAAATTCAGAAACAGCAATGCTGCGACAGACTGGGTGATTATTCACC 2880

Qy 958 AsnPheCysLeuPhePheTrpLysSerH1sGlnAspAsnH1sProLeuAlaSerLeuPro 977

Db 2881 AACTCTGCTGCTGTCTTCTTACAAATTCACACAGACAGTACCCCTTGGCCAGCTGCGCT 2940

Qy 978 LeuLeuGlyTrpSerLeuThrIleProSerGlySerGluAanIleGlnLysAspTrpVal 997

Db 2941 CTGCTGGGCTACTCCCTCAACATCCCTCTGAGTGTGAGAACATTCACAGAGACTATGTGG 3000

Qy 998 PheLysLeuAlaPheLysSerH1sValTrpTrpPheArgAlaGluSerGlyTrpThrPhe 1017

Db 3001 TTCAAGCTGCATTTCAAGTCCCAACCTGATCTACTTCAGGGCTGAAAGCAATACATTT 3060

Qy 1018 GluArgTrpMetGluValIleArgSerAlaThrSerSerAlaSerArgProH1sValLeu 1037

Db 3061 GAAAGGTGAGTGAAGTATCGAAGTGCACACAGCTCAGCTTCGGGCGCCACATCTTGG 3120

Qy 1038 SerH1sGlyGluSerLeuValTrp 1045

Db 3121 AGTCACAAAGATCTCACCCTGTAC 3144

RESULT 5
AKI17173
LOCUS
DEFINITION
AKI17173 3743 bp mRNA linear HTC 21-SEP-2005
Mus musculus activated spleen cDNA, RIKEN full-length enriched library, clone:FE30006D05 product:FERM, Rhodex, pleckstrin domain protein 2, full insert sequence.

ACCESSION
AKI17173
AKI17173.1 GI:74152707
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
PUBMED
10349636

REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
PUBMED
11042159

REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuaki, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasai, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Marahiki, M., Kameda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system-384-format
Genome Res. 10 (11), 1757-1771 (2000)
PUBMED
11076861

REFERENCE
AUTHORS
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Iehi, Y., Aizawa, K., Iwama, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Iwama, A., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Asahimura, M., Batalev, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nakado, I., Pezold, G., Quackenbush, J., Schriml, L.M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Baren, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., But, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustinich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombere, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, T., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.

CONSRM
TITLE
RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium
JOURNAL
NATURE 409 (6821), 685-690 (2001)
PUBMED
11217851

REFERENCE
AUTHORS
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., But, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalev, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalia, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Gustinich, S., Hirokawa, N., Jackson, I.D., Jarvis, E.D., Kanay, A., Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurouchin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numa, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.V., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlested, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavoian, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Maki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Iehi, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

CONSRM
TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
NATURE 420 (6915), 563-573 (2002)
PUBMED
12466851

REFERENCE
AUTHORS
Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Fritsch, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shinkawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Imbondato, A., Apweiler, R., Attaliya, R.N., Bailey, J.L., Banat, M., Baxter, L., Beisel, K.W., Beresno, T., Bono, H., Chalk, A.M., Chu, K.P., Choudhary, V., Christofides, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Furuki, S., Gariboldi, M., Georgi-Hemman, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustinich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ikey, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H.,

Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Lium, J., McWilliam, S., Madan Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavest, G., Pezole, G., Petrovsky, N., Piazza, S., Reed, J.F., Reid, J.F., Ring, B.Z., Ringwald, M., Roat, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Sempke, C.A., Sene, S., Seena, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sincalir, B., Sperling, S., Stupka, E., Sugita, K., Sultana, R., Takekura, Y., Taki, K., Tanno, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yeig, K., Yamanihi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Tesdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kato, T., Kawai, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Pleassy, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watanabe, A., Okamura-Ono, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y. FANTOM Consortium
The transcriptional landscape of the mammalian genome
Science 309 (5740), 1559-1563 (2005)

CONSRM
JOURNAL
PUBMED
REFERENCE
AUTHORS

7
Katsuyama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Firth, M., Ravasi, L., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.
Riken Genome Exploration Research Group
Antisense transcription in the mammalian transcriptome
Science 309 (5740), 1564-1566 (2005)

CONSRM
JOURNAL
PUBMED
REFERENCE
AUTHORS

8 (bases 1 to 3743)
Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Horii, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (14-APR-2004) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehito-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES
SOURCE

1. 3743
/organism="Mus musculus"
/mol_type="mRNA"
/strain="M01"
/db_xref="FANTOM_DB:F830006D05"
/db_xref="taxon:10090"

CDS

/clone="F830006D05"
/cbase_type="activated spleen"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
136. 3333
/note="unnamed protein product; PERM, RhoGAP and pleckstrin domain protein 2 (MGD) [MG:2385126 GB|BC009153, evidence: BLASTN, 99%, match=3696]
putative"
/protein_id="BAE42626.1"
/db_xref="GI:74152708"
R1RKLKIDSTVELFDIEPDCDGVLLTQWKNLLECDYFGLFKNVOSYWLPEM
KPIIRKPKNAVLAVKFPDPGQLOEYTRFLPALORDLLEERTCTANTA
ALLISHLQSEIDYDELDREHLKANEYLPNOKSELEKLDHQRHTQTPAESPO
VLEIARKLEMYGRFMAADREGTKNLINLVSHMGVVFQGTITINFMKSLSPK
RKRFLIKHBNVHGPIODTLPELFGSDECKNWKICVEHTTFRSLDQKPAKLVF
FSKSSFRYSGRTOQVLDVYKDGKMKRIPYERHSHKRTSLHALTVLPKQSVSFTD

Alignment Scores:

Pred. No.:	5, 86e-241	Length:	3743
Score:	2916.00	Matches:	562
Percent Similarity:	70.1%	Conservative:	179
Best Local Similarity:	53.2%	Mismatches:	268
Query Match:	53.4%	Indels:	48
DB:	6	Gaps:	9

US-09-555-342B-2 (1-1045) x AKT17173 (1-3743)

QY	1	MetGlyGluIleGlu-----GlnAyrProThrProGlySerArgLeuGlyAla	16
Db	136	ATGGAGAGATAGAGAAACATACAGAGCTTTTCCAACTTCAGAGACACCTGGGGTGC	195
QY	17	ProGluAnsSerGlyIleSerThrLeuGluArgGlyGlnIlySerProProThrProSer	36
Db	196	CAGACGCCATTGAGAGTACCTGAGACCAGACAGAGCTGTCCACAGATGCAG	235
QY	37	GlyIlySerLeuValSerIleIleGlnMetLeuAspAspThrGlnGlnIleAphGluVal	56
Db	256	GAGAGACATGCGGCATCAGAGTGAAGCTGTGACGACCTGTAGGATTATTCATC	315
QY	57	ProGlnArgAlaProGlyIlyValLeuLeuAspAlaValCysAsnHsIleAsnLeuVal	76
Db	316	GAGCCAAAGTATGACGGCGGCTTACCTGACTCAAGTGTGGAAGCATTTTGAATTCAT	375
QY	77	GluGlyAspTyrPheGlyLeuGluPheProAspHsIlyIleValIleThrValIleTyrLeu	96
Db	376	GAATGTACTACTTTTGTCTGGAGTTCAAGATGTCAGTCCATCGATTGGCTTGA	435
QY	97	LeuLeuIlySerIleValIlyGlnIleArgArgProIlyHsIleValIlyValIlyPheVal	116
Db	436	CGATGAAACCATCATCTTAGGCAGATGACGAAACCAAAAATGCGGTTCGCTCGCA	495
QY	117	ValIlyPhePheProProAspHsIleThrGlnLeuGlnGluIleuLeuThrArgTyrLeuPhe	136
Db	496	GTAATAATTTTCCCGCTGACCTGTGTGACGTGCAAGAAAGATACCAAGTACCTGTTT	555
QY	137	AlaLeuGlnValIlyGlnAspLeuAlaGlnGlyIlyArgLeuThrCysAsnAspThrSerAla	156
Db	556	GCTTGCATCTCAAGAGACCTCTCGAGAAACGCTTGACTGACGCGCAACATCGCA	615
QY	157	AlaLeuLeuIleSerHsIleValIlyGlnSerGluIleGlyAspPheAspGluAlaLeuAsp	176
Db	616	GCCCTTCTATATCCACCTTGTGAGTGGAAATGAGATATATGATGAACCTTGAT	675
QY	177	ArgGlnHsIleuAlaIlyAsnIlyTyrIleProGlnGlnAspAlaLeuGluIlyAspIle	196
Db	676	CGAGAACCTCTCAAGACCATGATGATCTTGCACAGAGGAATCTTCAAGAAAGATA	735
QY	197	ValGluPheHsIleAsnHsIleGlyGlnIleThrProAlaGluIlySerAspPheGlnLeu	216
Db	736	CTAGACTTCATAGAGGACACGCGGCAGACTCTGCAAGTACGATTTCCAGGTCTT	795

QY 217 GluIleAlaArgXLeuGluIleuMetTyrGlyIleArgLeuHisProAlaLysAspArgGlu 236
 Db 796 GAGATTGCAAGAGAACTGGAAATGTAAGCAATGCGATTGCGATGGCTTTCAGAGAA 855
 QY 237 GlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
 Db 856 GGGACCAAGATTAACTAGCAGATTCTCCATGGGTGTCTGTCTTCAGGGATCCACC 915
 QY 257 LysIleAsnAlaPheAsnThrAlaLysValAlaGlyLysLeuSerPheLysArgLysArgPhe 276
 Db 916 AAAACCAATACCTTAACCTGTCGAAAGCTCCAAAGCTTAAGCTTCAAGAAAGGTTT 975
 QY 277 LeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
 Db 976 CTTATCAAACTCCACCTGAGGTCCATGGGCTTCACAGACACCTTAGAGTTCTTGTG 1035
 QY 297 AlaSerArgAspPheCysLysSerPheThrLysIleCysValGlnHisHisAlaPhePhe 316
 Db 1036 GGTAGCAGAGATGAATGTAAGAACTTTCGAAAGATATGTGTGAGTACATACCTTTT 1095
 QY 317 ArgLeuPheGluGluProLysProLysProLysProLysProLysPheSerArgLysSer 336
 Db 1096 AGACTCTGTGACCAAGCTTAAGCCAAAGGCAAGGCTGCTTCTTCAGCCAGGCTCTCC 1155
 QY 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyLysHis 356
 Db 1156 TTCAGATTCAGTGGAAAGAACTCAAGAAACAATAGATTAATGTAAGACGGTGGAAATG 1215
 QY 357 LysLysValGlnPheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSer 376
 Db 1216 AAGAGAATTCATAGAAAGACGACAGTAAAGACTCGG----- 1254
 QY 377 GlnProThrGluLeuAsnSerGluValLeuGluGlnSerGlnGlnSerThrSerLeuThr 396
 Db 1355 -----ACATCTCTTCACTGCTGTGACTGTAGATCGCTTAACAGAGGCTCTTCAAC 1308
 QY 397 PheGlyGluGlyAlaGluSerProGlyGlnGlnSerCysArgArgGlyLysGluProLys 416
 Db 1309 GATGGCTTGAGAACTTCTGCC-----TCC 1332
 QY 417 ValSerAlaGlyGluProGlySerHisProSerProAlaProArgArgSerProAlaGly 436
 Db 1333 CTGTCTTCAGCAAAATGTCTCTTATATCAACCCCTAGTCTTCTTGTGCTCCCTCGT 1392
 QY 437 AsnLysGlnAlaAspGlyAlaAlaSerAlaProThrGluGluGlnGluValLys 456
 Db 1393 CTGCCCAATTGTAAGACAGCAGCAGCTCTTGTGATCCGACGCTCCGTCATCAAG 1452
 QY 457 Asp-----ArgThrGlnGln 461
 Db 1453 AGCACACGACGAGAGAGACGCGACCATCATCATCTGACGCGCCCAACACACAGTCCG 1512
 QY 462 SerLys---ProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
 Db 1513 GCCCATCTCCCTGAGACCCCGGTGCTCCGGCTGTCCAGGCTTTTCATGAGATGTCT 1572
 QY 479 HisLeuSerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeu 498
 Db 1573 CAGCCTTCTCCCTCAGGCTGTAAGCCACCTGACCTGTGCTCT-----GAGCTT 1623
 QY 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuLeuAsn 518
 Db 1624 CAGCGCGGCTGAGTACAGCTGACGAGGTCATCCCTGTGCTACGCGCTGTGTCTAGT 1683
 QY 519 AspGlnAlaCysProArgThrAspArgLysAspGluGlyArgArgLysArgPheProThr 538
 Db 1684 GCGCGTGTGACGCGCAGATGGAACAACAAAGACAGAAACACACACATCCGACAGA 1743
 QY 539 AspLysAlaIleTyrPheIleAlaLysGluValSerThrThrGlnArgThrThrLysLysAsp 558
 Db 1744 GATGAGGCTTAATTCATAGCCCAAGAGATTCCTCGTACAGAAACCACTATCTGGAAGAT 1803
 QY 559 LeuGluValIleThrSerThrPheGlnSerThrValSerLysGluAspAlaMetProGlu 578

Db 1804 TTAGAGATTATCACTGTGTGCTTCCGACGCTCGATCAAGAGAGGACCATCGCTGCA 1863
 QY 579 AlaLeuLysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAsnPhe 598
 Db 1864 GCCCTGATGCGCTGCTTTTCTCAACATTAATCAGCTACAGAGTTCCACAGAGGCTTC 1923
 QY 599 LeuLysGluIleGluGlnArgLeuAlaLeuThrGluGlyAspSerAsnAlaGlnIleArg 618
 Db 1924 CTTACAGAGGTGAACAGAGCTGACCTGGGAAGGCGCTCCAGTGGCCACTTAAGA 1983
 QY 619 ---AspTyrGlnArgGlyLeuAspValMetLeuLysAsnIleGlnLysMetLysHisLeu 637
 Db 1984 GGTATTCACAGCGAATCGGGGACATCTCTCCAGAAACATCGCTTAAGGAATTT 2043
 QY 638 AlaAlaHisLeuThrLysHisSerGluAlaLeuGluAlaLeuGlnGlyLysSer 657
 Db 2044 ACTAGCTACTTCCAAAGACAGATGAGGCTCTTAACAGAACTGAAAGGCCACAAACAC 2103
 QY 658 SerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuPro 677
 Db 2104 TGTAAAGCTGAGAGGACGTCTAACAAAGTTTGAAGTTCGAAAGGTGTGCTACCTGCT 2163
 QY 678 LeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisTyrLysGlnValLeuGlu 697
 Db 2164 CTCACACATTCCTGCTGAAACCCGTCACAGGCTAGTCCACTACCTGTGCTGAGC 2223
 QY 698 ArgLeuCysLysHisHisProProSerHisAlaAspPheArgAspCysArgAlaLeu 717
 Db 2224 CGCGTGTGTCTCACTCTCTCTGCGCACCGCACTATGCGGCTGCAAGAGGCACTG 2283
 QY 718 AlaGluIleThrGluMetValAlaGlnLeuHisGlyHisMetIleLysMetGluAspPhe 737
 Db 2284 AAGGCATCACAGAAAGTACCAACGAGTCCAGCAAGGCTTACCCGCTGAAACCTA 2343
 QY 738 GlnLysLeuHisGluLeuLysLysAspLeuIleGlyIleAspAsnLeuValProGly 757
 Db 2344 CAGAAATTGACGAGAGCTACAGCAGAGACTCGTGGTGAAGAACTCATTTGCTCGG 2403
 QY 758 ArgGluPheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnArg 777
 Db 2404 AGGAGATTATCCGTGAGGCTGCTGCAACAGCTCAACAAAGAGGCTTCCACAGAGG 2463
 QY 778 MetPhePheLeuPheAsnAspValLeuLeuThrThrSerArgGlyLeuThrAlaSerAsn 797
 Db 2464 ATGTTTTCTGTTCTCAGATATGTACTGTATCAAGCAAAAGTGTACAGAGCCAGT 2523
 QY 798 GlnPheLysValHisGlyGlnLeuProLeuTyrGlyMetThrIleGluGlnSerGluAsp 817
 Db 2524 CATTTCCGAGATCCGTGCTTCTTCCACTCCGTGGCATCGTGAAGAAAGTACAGAT 2583
 QY 818 GluTyrGlyValProHisCysLeuThrLeuArgGlyLysArgGlnSerIleIleValAla 837
 Db 2584 GAATGTGTGTTCTTCAATTGCTTCACTATACAGTACAGTCAAGAAACATATGTGTGACA 2643
 QY 838 AlaSerSerArgSerGluMetGluLysTyrValAlaLysArgIleGlnMetAlaIleAspLeu 857
 Db 2644 GCCAGCACTCGCTAAGAAAGAAAGTGAAGAGAGCTGAATGCAGCAATCAA--- 2700
 QY 858 AlaGluLysSerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLys 877
 Db 2701 GCAGCCAAAGACTATTCGGATCTACCCCAATGCTGTGGAGGCGCGGTATATCTCT 2760
 QY 878 SerProAspGluAlaThrAlaAlaAspGlnLysSerGluAspAspLeuSerAlaSerArg 897
 Db 2761 ACCCTGATGATTTCTGATGAAGTCTCTGTAAGAAATCAGAAAGTGTGAGAGAAACCG 2820
 QY 898 ThrSerLeuGluArgGlnAlaProHisArgGlyAsnThrMetValHisValCysThrHis 917
 Db 2821 GGTCTCCGTGAGGAAACGCGACAGACCGGGCCAAATACACAAATGCAATGTGTGCTGATC 2880
 QY 918 ArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerGly 937

Db 2881 CGTAAATCAAGTGTGTCCAGACGACACAGTCAGCTGTGGAAACGACTTTCAGGA 2940

Qy 938 AsnLeuLeuArglySerPheLyAsnSerAsnGlyTTPGlnLyLeuTPValPheThr 957

Db 2941 TATCTGCTGAGAAAGTTCAAGAACAGTAATGCTGGAGAAAGCTCTGGGTGCTTTAAC 3000

Qy 958 AsnBheCysLeuPhePheTyTyLySerHisGlnAspAsnHisProLeuAlaSerLeuPro 977

Db 3001 AACTTCGCTGTTGCTTCTTAAACATCATCAGATGACTACCCCTGGCAGGCTCCCA 3060

Qy 978 LeuLeuGlyTyTySerLeuThrHisProSerGluSerGluLeuAsnHisGlnLyAspTyTyVal 997

Db 3061 CTATTGGGTTACACCGGAGGCTCCCGAGGAGGCTACACATCCCAAAAGACTAATGTC 3120

Qy 998 PheLyLeuHisPheLySerHisLeuValTyTyTyPheArgAlaGluSerGluTyTyPhe 1017

Db 3121 TTCAAGCTCCAAATTCAAATCTCAGCTTACTTCTTCGGGCTAGAGCAAGTACACATTT 3180

Qy 1018 GluArgTTPMetGluValHisSerAlaThrSerSerAlaSerArgPro 1034

Db 3181 GAAAGGTGATGACGTCATCAAAAGGCGCAGCTCACAGGAGGACCC 3231

RESULT 6

AKO50860 3874 bp mRNA linear HTC 02-SEP-2005

LOCUS AKO50860.1 GI:26094164

DEFINITION Mus musculus 9 days embryo whole body cDNA, RIKEN full-length enriched library, clone:DO30026M03 product:weakly similar to CDEP (Homo sapiens), full insert sequence.

ACCESSION AKO50860

VERSION AKO50860.1 GI:26094164

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE 1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 10349636

REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 11042155

REFERENCE 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kikunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Miehine, T., Harada, K., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishiji, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Onari, E., Matsubara, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, D., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 11076861

REFERENCE 4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

REFERENCE 5 The FANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

REFERENCE 6 RIKEN Genome Exploration Research Group, Genome Science Group

TITLE (Genome Network Core Team) and the FANTOM Consortium. Antisense transcription in the Mammalian Transcriptome Science 309, 1564-1566 (2005)

REFERENCE 7 The FANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group). The transcriptional landscape of the Mammalian Genome Science 309, 1559-1563 (2005)

REFERENCE 8 (bases 1 to 3874)

ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Koda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

Direct Submission

TITLE Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

LOCATION/Qualifiers

1. 3874

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:DO30026M03"

/db_xref="taxon:10090"

/clone="DO30026M03"

/issue_type="whole body"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="9 days embryo"

1. 3874

/note="putative weakly similar to CDEP (Homo sapiens) (SPTR:Q9Y4F1, evidence: PASTY, 53.7%ID, 99.7%length, match=1125)"

ORIGIN

Alignment Scores:

Pred. No.: 3,03e-239 Length: 3874

Score: 2896.50 Matches: 565

Percent Similarity: 70.3% Conservative: 181

Best Local Similarity: 53.3% Mismatches: 261

Query Match: 53.0% Indels: 55

DB: 6 Gaps: 12

US-09-555-342B-2 (1-1045) x AKO50860 (1-3874)

Qy 1 MetGlyGlnHisGlu-----GlnArgProThr-ProGlySerArgLeuGlyAla 16

Db 70 ATGGAGAGAGTAAAGCAACATACAGAGCTTTACCACTTCAGGAGACCGCTTGGGTGCC 129

Qy 17 ProGluAsnSerGlyLeuSerThrLeuGluArgGlyGlnLyfProProProSer 36

Db 130 CACACCGCATTCGATGAGTACCTCGAGGACGACAGAGCGCTGCACCCAGAGTGCAG 189

Qy 37 GlyLyLeuValSerLeuLyHisGlnMetLeuAspAspThrGlnGluAlaPheGluVal 56

Db 190 GAGAAAGCATGGCGATCAGAGTGAAGCTGGACACACCTGTAGATTTTACATC 249
Qy 57 ProGlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuVal 76
Db 250 GAGCGCAAGTGTGACGGGACAGGTCTTACTGACTCAAGTGTGGAAGCATTTTGAAATCTCAAT 309
Qy 77 GlnGluAspTyrPheGlyLysGluPheProAspHisLysLysIleThrValTyrLeuAsp 96
Db 310 GAATGTGACTACTTGTGTGAGTTCAAGAAATGTCCAGTCTTACTGGAATTTGGCTTAA 369
Qy 97 LeuLeuLysProIleValIleGlnIleArgArgProLysHisValValIleLysPheVal 116
Db 370 CCTATGAAGAACCATATTAGCGAAGTACGAAGAACCAAAATATGCGGTCTTCGCTGGCA 429
Qy 117 ValLysPhePheProProAspHisThrGlnLeuGlnGluLysLeuThrArgTyrLeuPhe 136
Db 430 GTAAATATTTTCCGGCTGACCTGGTCAAGTGTGCAAGAAAGATACAAAGGCACCTGTT 489
Qy 137 AlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAla 156
Db 490 GCCTTGCACTCAAGAGACCTCTGGAGAAAGCTGACCTGACCGGCCAACACTGCA 549
Qy 157 AlaLeuLeuLysSerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAsp 176
Db 550 GCCCTTCTCATATCCCACTTCTGCACTCGGAAATCGAGATTTATGATGAAGAACCTTGAT 609
Qy 177 ArgGluHisLeuAlaLysAsnLysTyrIleProGlnIleAspAlaLeuGluAspLysIle 196
Db 610 CGAAGAACCTCAAGCCAAATGAATPACTTGCCCAACGAGAAATCTCTGAAGAAAGATA 669
Qy 197 ValGluPheHisLeuAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuLeu 216
Db 670 CTAACTTCCATCAAGGACACCGGGCCAGACTCCCGAGAGTCCGATTTCCAGGTGCTT 729
Qy 217 GlnIleAlaArgArgLeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGlu 236
Db 730 GAGATTGCAAGAGACTGGAATGTATGGCATACAGTTTCCACATGGCTTTCAGACAGAA 789
Qy 237 GlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
Db 790 GGGACCAAGATTAACTGTAGCAGTTTCTCACTGGGTGTCTGTGTGTCCAGGGTACACC 849
Qy 257 LysIleAsnAlaPheAsnThrAlaLysValArgLysLeuSerPheLysArgLysPhe 276
Db 850 AAAATCATATACCTTCAACTGTCCAAAGTCCGTAACCTTAACGTAAGAAAAAGTTT 909
Qy 277 LeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
Db 910 CTATCAAACTCCACCTGAGAGTCCATGGCCCTTACCGACACGCTTGAAGTTCTTGTG 969
Qy 297 AlaSerArgAspPheCysLysSerPheTyrLysIleCysValGlnHisAlaPhePhe 316
Db 970 GGTAGCAGAGATGATGATGAACCTTCTGAAAGATATGTGTGAGTACCATACCTTTT 1029
Qy 317 ArgLeuPheGluGluPheLysProLysProLysProValLeuPheSerArgGlySerSer 336
Db 1030 AGACTCTCTGACAGCCCTAAGCCAAAGGCAAGGCTGTCTTCTCAAGCCGAGGCTCC 1089
Qy 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGlnGlyHis 356
Db 1090 TTCAAGTACACTGGAAGAACTACAGAAACACTAGAGATTATGTCAAGAGCGTGGAAATG 1149
Qy 357 LysLysValGlnPheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSer 376
Db 1150 AAGGAATTCATCGAAAGACGCGCACAGTAAGATCGG----- 1188
Qy 377 GlnProThrGluLeuAsnSerGluValLeuGlnGlnSerGlnGlnSerThrSerLeuThr 396
Db 1189 -----ACATCTTCTTACGTCTGACTGATGATCTGCTTAAACAGAGGCTCTCCCTCAC 1242
Qy 397 PheGlyGlnGluAlaGluSerProGlyLysGlnSerCysArgArgGlyLysGluProLys 416

Db 1243 GATGGCTTGAAGACTTCTGC-----TCC 1266
Qy 417 ValSerAlaGlyLysProGlySerHisProSerProAlaProArgSerProAlaGly 436
Db 1267 CTGTCTTCAAGAAATGTCTCTTTATTCACCCCTGAGTTCTTGTGTCCCTCGGT 1326
Qy 437 AsnLysGlnAlaAspGlyAlaAlaSerAlaProThrGlnGlnGlnGluValValLys 456
Db 1327 CTGCCCAATTGGAAGACAGACAGACGCTCTTGTGATCCGACGCTCCCGTATCAAG 1386
Qy 457 Asp-----ArgThrGlnGln 461
Db 1387 AGCAGACAGACAGAGAGACGCGACCATCATCATCTGACGGCCCGACAGACAGTGG 1446
Qy 462 SerLys---ProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
Db 1447 GCCCATCTCCCTGAGACCCCGGTGCTCGGCTGGTCCAGGCTTTTCTAGATATGTCCT 1506
Qy 479 HisLeuSerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeu 498
Db 1507 CAGCTTCTCCCTCCAGCTTAAGACCCCTGACCTGTGCTC-----GAGCTT 1557
Qy 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuLysSerProLeuAsn 518
Db 1558 CAGCGCGGTGAGTACAGCTGACGACGAGGTTCATCCCTGTGCTCAGGCTGTGCTCAGT 1617
Qy 519 AspGlnAlaCysProArgThrAspAspGluAspGluArgArgLysArgPheProThr 538
Db 1618 GGCCTGTGTACAGCAGAGATGGAACACCAAGAAAGCAAGAACACATGCCAGAA 1677
Qy 539 AspLysAlaTyrPheIleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAsp 558
Db 1678 GATAGGCTTATTTCAATGACCAAGAGATTTCCCTGACAGACCACTATCTGAAGAT 1737
Qy 559 LeuGlnValIleThrSerTyrPheGlnSerThrValSerLysGluAspAlaMetProGlu 578
Db 1738 TTAAAGATTATCACTGTGTGTTCCGGAGCTGTGATCAAGAGAGAGGCATCTGCA 1797
Qy 579 AlaLeuLysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisIleAsnPhe 598
Db 1798 GCCCTGTGGCCCTGCTTTTCTCAACATTGATCAAGTTCACAGTTCCACAGAGGCTTC 1857
Qy 599 LeuLysGlnIleGlnGlnArgLysLeuAlaLeuTyrGluArgSerAlaGlnIleArg 618
Db 1858 CTCAACAGAGTGAACAGAGGCTGCACTCTGGAAAGGCCCTCAAGTCCCACTTAA 1917
Qy 619 ---AspTyrGlnArgIleGlyAspValMetLeuLysAsnIleGlnIleMetLysIleLeu 637
Db 1918 GGTATCACAGCGAATCGGGGACATCCTCTCGAAGAACATGGCTATTAAGAAATTT 1977
Qy 638 AlaAlaHisLeuTyrLysHisSerGlnAlaLeuGlnAlaLeuGlnAsnGlyIleLysSer 657
Db 1978 ACTGACTACTTCCAAACACAGATGAGTCTTAACAGAACTGGAAGGCGCCACAAACAC 2037
Qy 658 SerArgArgLeuGlnAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuPro 677
Db 2038 TGTAAAGCTGGAAGGACAGTCTACAAAGATTTAAGTCTCAAAAGCTGTGTACTGCT 2097
Qy 678 LeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisTyrLysGlnValLeuGln 697
Db 2098 CTCAACACATTCCTGTGAAGCCCGTCCAGAGGCTATGCTCCAGTCCGCTGCTGAGC 2157
Qy 698 ArgLeuCysLysHisIleProProSerHisAlaAspPheArgArgCysValArgAlaLeu 717
Db 2158 CGGCTGTGTCTACTCTCTCCGAGACCGGACGCACTATGCGGACTCCATGAGGCACTG 2217
Qy 718 AlaGluIleThrGlnMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPhe 737
Db 2218 AAGGCATCACAGAAAGTACACCGAGACTCCAGAAAGCCTTACCGCGCTGAAAACCTA 2277
Qy 738 GlnLysLeuHisGluLeuLysLysAspLeuIleGlyIleAspAsnLeuValValProGly 757
Db 2278 CAGAAATTGACGAGGCTACAGCGAGACCTGGTCCGTGAGAAACCTCATGTGCTCGGG 2337

```
Qy 758 ArgGluPheIleAArgLeuGlySerLeuSerIleuSerGlyIleuGlyLeuGlnArg 777
Db 2338 AGGAGATTATTCCTGAGGGCTGCTGCACAGCTCCACCAAGAGGGCTGCACACAGG 2397
Qy 778 MetPhePheIleuPheAenAapValIleuLeuTyThrSerArgIleuThrAlaSerAsn 797
Db 2398 ATGTTTCTCTGTTCTCAGATATGTTCTGTATACAGCAAAAGATGTCACAGAGCCAGT 2457
Qy 798 GlnPheIleuValIleuGlyGlnIleuProLeuTyThrIleuGlnIleuGlnIleuAsp 817
Db 2458 CATTTCGGAGTCCGTGCTTCCTCCACTCGTGGCAATGCTGTAGAGAAAGGAAAT 2517
Qy 818 GluTyPrlValProHisCysLeuThrLeuArgIleuGlnArgIleuSerIleuValAla 837
Db 2518 GAATGGCTGTCTCTCTCATGTTTCAACATCATACCTGACCTCAAGAAACATTTGCTGCA 2577
Qy 838 AlaSerSerArgSerGlnMetGluIleuTyPrlValGluIleuPrlIleuMetAlaIleuAsp 857
Db 2578 GCCAGCACTCGGCTAGAAAGAAAGATGATGACAGACCTGAATGCAGCATCCAA--- 2634
Qy 858 AlaGluIleuSerSerSerProAlaProGluPheIleuAlaSerSerProProIleuAsn 877
Db 2635 GCAGCCAAAGACTATCGGTGACTCAGCCCACTGCTGTGGAGGCCGCTGTATATCTGT 2694
Qy 878 SerPro-----AapGluAlaThrAlaAlaIleuGlnIleuSerGluIleuAspIleuSer 894
Db 2695 ACCCTTGATCTTGATGAAAGTCTCTG-----GAAGATCAGAAAGATGCTCG-AGG 2747
Qy 895 AlaSerArgThrSerLeuGlnIleuArgGlnAlaProHisArgGly--AanThrMetValHis 913
Db 2748 AAAGCCGGGGCTGCCCTGAGAGGGAGAACAGCCAGCCAGGGCCCAATATACAAATGCAT 2807
Qy 914 ValCysThrPrlIleuArgIleuThrSerValSerMetValIleuPheSerIleuAlaValGluAsn 933
Db 2808 GTGGCTGTGATCCGATATACAGATGTGTCCAGACCAACCACTGACGCTGTGAGAAC 2867
Qy 934 GlnIleuSerGlnIleuLeuIleuArgIleuPheIleuAsnSerArgIleuTyPrlIleuLeuTyPrl 953
Db 2868 CAGTTTCAGAGATATCTGCTGAGAAAGTTCAAGAACAGTAAAGTCTGCGACAGAGCTCGG 2927
Qy 954 ValValIleuThrIleuPheCysLeuPheIleuTyThrSerHisGlnIleuAsnIleuProLeu 973
Db 2928 GTGGCTTTTACCAACTCTGCTGTCTCTTAATAACATCAGATGATGATACCCCTG 2987
Qy 974 AlaSerIleuProLeuIleuGlyTyThrSerLeuThrIleuProSerGlnIleuAsnIleuGln 993
Db 2988 GCCAGCTCCACATATGTTGGTTACAGGCTGAGCTCCCAAGGAGGCTGACAGATCCAC 3047
Qy 994 IlyAaPrlTyValPheIleuValIleuPheIleuSerHisValTyThrPheArgAlaGluSer 1013
Db 3048 AAAGACTATGTTCTTCAAGCTCCCAATTCATCTCACTTCTTCCGGGCTGAGAGC 3107
Qy 1014 GluTyThrPheGlnIleuArgIleuPheGluIleuValIleuSerArgIleuThrSerSerAlaSerArg 1033
Db 3108 AAGTACACATTTGAAAGTGTGATGACATCATCAAAAGGCCAGACACTCACCGAGGAGA 3167
Qy 1034 Pro 1034
Db 3168 CCC 3170

RESULT 7
AY415515 3165 bp DNA linear GSS 17-DEC-2003
LOCUS DEFINITION Homo sapiens FARP2 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY415515
VERSION AY415515.1 GI:39771474
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
REFERENCE Homidae; Homo.
1 (bases 1 to 3165)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3165)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
location/Qualifiers
1..3165
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>3165
/gene="FARP2"
/locus_tag="HCMS583"
gene
ORIGIN
Alignment Scores:
Pred. No.: 2,92e-189 Length: 3165
Score: 2318.50 Matches: 486
Percent Similarity: 59.9% Conservative: 149
Best Local Similarity: 45.8% Mismatches: 368
Query Match: 42.4% Indels: 57
DB: 14 Gaps: 14
US-09-555-342B-2 (1-1045) x AY415515 (1-3165)
Qy 1 MercGlyGluIleGlnGlnArg-----ProThrProGlySerArgLeuGlnValAla 16
Db 1 ATGGGGAGATTAAGAAACATACAGAGTCTGACAGCTGACAGGAGTGGCTTGGGGTCC 60
Qy 17 ProGluAsnSerGlyIleuSerThrLeuGlnIleuArgGlyGlnIleuAspProProThrProSer 36
Db 61 CAGACCCCTGTGGAGATTACAGCCCTTGAGCTGGGCAAGCTCTTGGCCAGAAATGCAA 120
Qy 37 GilyLeuIleuValSerIleuValIleuMetLeuAspAspThrGlnIleuAlaPheGluVal 56
Db 121 GAGAAAGCACTGCACCTCAGAGTAAAGCTGCTGACAAACCATGGAATATTTGACATT 180
Qy 57 ProGlnArgAlaProGlyIleuValIleuLeuAspAlaValCysAsnHisIleuAsnIleuVal 76
Db 181 GAGCCTTAATGCGATGGCCAGGATTAATGACAAAGTGTGGAAGCGTTTAAACCTGTGA 240
Qy 77 GluGlyAspTyPrlPheGlyLeuGlnIleuPheProAspHisIleuValSerIleuThrValIleuAsp 96
Db 241 GAATGTGACTTACTTCGGGATGAGATTTCAAATATCTCAAGTCCATCTGAGATTTGGCTTGA 300
Qy 97 LeuIleuAspProIleuValIleuGlnIleuArgArgProIleuHisValValValIleuPheVal 116
Db 301 CCTATGAAACCATCATTTAGGCAATATACGAAGCCAAAGAAATGTGGTGGCTTGCCTAGCT 360
Qy 117 ValIleuPhePheProProAspHisIleuGlnIleuGlnGluIleuLeuThrArgTyIleuPhe 136
Db 361 GTAAATTTTTCACACTGATCTGTGTCAGCTACCAAGAAAGAAATATCAAAATCTTGT 420
Qy 137 AlaLeuGlnValIleuGlnIleuAspIleuAlaGlnGlyArgLeuThrCysAsnAspThrSerAla 156
Db 421 GCCCTGCACTTAAAGAGACCTCTGAGAGACCGTTTGACTGTGTGACACACACAGCG 480
Qy 157 AlaLeuIleuIleuSerHisIleuValGlnIleuSerGluIleuGlyAspPheAspGluAlaLeuAsp 176
```


Db	481	GCCTCTTCACGCTCCCTCATCTCTCGAGTCGGAATAAGAGATTACGAATGAACGCTGAC	540
Oy	177	ArgGlnH1eLeuAlaLeuAenlyeThrlleProGlnGlnAerAlaLeuGluAerlylle	196
Db	541	CGAGAGCACTTAAGTGAAGCAAGTATTGCTGTGGCCAGACACTGCTTGAAGAAATA	600
Oy	197	ValGluPheH1eH1eAenH1eIleGlyGlnThrProAlaGluSerAerPheGlnLeuLeu	216
Db	601	CTAGAATTCCATCAGAAAGCAGTGGGGCCAGACACTGCTGAATCGGATTCCAGGTCCTC	660
Oy	217	GluIleAlaAerGluLeuGluMetGlyGlyIleAerGluH1eProAlaLeuAerArgGlu	236
Db	661	GAATATGCTCGAAATGTGGAAATGTAACGGCAATTCAGATTCAGCTTCTGACAGGGA	720
Oy	237	GlyThrLyS1eAenLeuAlaValAlaAenThrGlyIleLeuValPheGlnGlyTherHr	256
Db	721	GGAACTCAAGATTCAATGCGAGTTCCTCCACATGGGTGATCTGTCTTCCAGNNNNNNNN	780
Oy	257	LyS1eAenAlaPheAenThrAlaLeuValAerGlyLeuSerPheLyAerGlyAerPhe	276
Db	781	NN	840
Oy	277	LeuIleLyLeuAerProAerAlaAenSerAlaIyGlnAerThrLeuGluPheLeuMet	296
Db	841	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNGACCTTCCAGACACATTTGAATTTTGTTG	900
Oy	297	AlaSerArgAerPheCyAlySerPheThrlLyS1eCyValGlnH1eAlaPhePhe	316
Db	901	GGTATTAAGATGAATGTAAAGACTTCGGAAGATTGTGTGAATATCAACTTTT	960
Oy	317	ArgLeuPheGluGluProLyAerProLyAerProValLeuPheSerArgGlySerSer	336
Db	961	AGACTTTTGGACCAACTTAAGCCAAAGAAAGCGCTTCTTCAAGCGGGGCTCTCC	1020
Oy	337	PheArgPheSerGlyAerThrGlnLyGlnValLeuAerGlyValLyGluGlyGlyH1s	356
Db	1021	TTCAATATCAGNN	1080
Oy	357	LyS1eValGlnPheGluAerGlyH1eSerLyS1eH1s---SerIleAerSerLeu---	374
Db	1081	NN	1140
Oy	375	AlaSerGlnProThrGluLeuAenSerGlyValLeuGluGlnSerGlnInserTherSer	394
Db	1141	CGAGACCTACCAAAACAG-----AGC	1160
Oy	395	LeuThrPheGlyGluGlyAlaGluSerProGlyGlyGlnSerCyAerArgHr-----	411
Db	1162	ATCTCATTTCCCGAGAGGATGAGACTCCTGCTCCCATCTTCAGAGATGCTTTTAC	1220
Oy	412	-----GlyLyGluProLyAerAlaGlyGluProGly-----	423
Db	1222	TCGCTCTCTCCCTCACTCTGTGCTCCCTGTGGCTGCGCAGATTAAAGACAGCAGAC	1280
Oy	424	-----SerH1eProSerProAlaProAerArgSerProAlaGlyAenlyeGlnAlaAer	441
Db	1282	TCCCTCACAGATCCCGAGTTCTTCACTCAAGATCCAGCTGCGACAGAGCGCGTGA	1340
Oy	442	GlyAlaAlaSerAlaProThrGlnGluGluGluGluValAlaLyAerArgThrGlnGln	461
Db	1342	GCAGTGGCTGGAGGCCCGAC-----ACACCATCG	1370
Oy	462	SerLyAerProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro	478
Db	1372	GGCCAGCCCCCTGGGGCCCCGCACTCAAGCTGTGTCCAGGCGCTTCCAGGAAGTCTT	1430
Oy	479	H1eAerSerGluLeuSerValAenSerGlnGlyValAlaProAlaAenValThrLeu	498
Db	1432	CAGCTTCTCCCTCCAGCCGGAAGACCCCCGAGCTGAGCCCTGCA-----TTT	1480
Oy	499	SerProAenLeuSerProAerThrLyGlnAlaSerProLeuIleSerProLeuAen	518

Db	1483	CAGGTCGCTTTGGGCGCCAGCTGAACAGGGGCTCATCCCACTCTCGAGCCCTGTCTCAGT	1542
Qy	519	AspGlnAlaCysProArgThrAspAspGluAspGluGlyIleArgArgIysArgPheProThr	538
Db	1543	GATGCTGGCGGAGCGGGATGGACTGGCAG--GAGCCACAGACAAGNNNNNNNNNNNN	1599
Qy	539	AspIlyValArgPheIleAlaLeuValSerThrArgIleValGlnArgIleLeuLysAsp	558
Db	1600	NN	1659
Qy	559	LeuGluValIleThrSerTrpPheGlnSerThrValSerIlyAspAlaMetProGlu	578
Db	1660	NN	1719
Qy	579	AlaLeuLysSerLeuIlePheProAspPheGluProLeuHisLysPheHisThrAspPhe	598
Db	1720	ACTGTGATGAGCGCTGCTCTTCTCCACATCATCATCCATCATATGAGTTCCACAGAGGCTTC	1779
Qy	599	LeuIlyGluIleGluGlnArgLeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIleArg	618
Db	1780	CTGGCGCGAGTGGAGCAGAGGCTGGCACTCTGGGAAGGCCCTCCAAAGCCCAACAATA	1839
Qy	619	---AspArgIleGlnArgIleGlyAspValMetLeuLysAsnIleGlnIlyMetLeuHisLeu	637
Db	1840	GCGAGCTACACAGATCGGGGACATCTGTGCTCGAGAACTGCGCCAGTTTAAAGAGATT	1899
Qy	638	AlaAlaHisLeuTrpLysHisSerGluValLeuGluValLeuGluAsnGlyIleLysSer	657
Db	1900	ACCGACTACTTCCAAAGACATGACAGAGGTCTTACAGAACTGGAAAAGGCTACCAAGCC	1959
Qy	658	SerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnIlyValCysTrpLeuPro	677
Db	1960	TGTAAAGAGTTGGAGGACAGTGTACAGAGAGTTTGAAGCTGCGAAGAGTCTGACTTGGCT	2019
Qy	678	LeuAsnTrpPheLeuLeuArgProLeuHisArgIleuMetHisTrpLysGlnValLeuGlu	697
Db	2020	CTCAACAGCTTCTCGTGTGAAGCCCATCCAGGGCGCTGCACTACCGCTGCTGCTGGC	2079
Qy	698	ArgLeuCysLysHisIleProProSerHisIleAlaAspPheArgAspCysArgAlaAlaLeu	717
Db	2080	CGCCTATCGGACATTAACAGCCCGCGGACCATATACGTGATCGTCAGTGAAGCCCTG	2139
Qy	718	AlaGluIleThrGluMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPhe	737
Db	2140	AAAGGCATCACAAGAGTGTACACACACACATACAGACATTCATCTCCGGCTGGAAACTGTG	2199
Qy	738	GlnIlyLysLeuHisGlyIleuLysLysAspLeuIleGlyIleAspAsnLeuValProGly	757
Db	2200	CAGAAAGCTAACCGAGCTGACCGGAGCTGTGTGGCATAGAAACCTCATTTGCTCTGGC	2259
Qy	758	ArgGluPheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnIlyArg	777
Db	2260	AGGAGATTCAATCGTGAAGGCTGCTTCAACAGCTCACCAAGAAAGGCGCTGCACAGAG	2319
Qy	778	MetPhePheLeuPheAsnAspValLeuLeuTrpThrSerArgGlyLeuThrAlaSerAsn	797
Db	2320	ATGTTTTTCTGTCTCAGATATATGTGTGTGTACCAAGCAAAAGAGTGTGCGAGACCGAG	2379
Qy	798	GlnPheLysValHisGlyGlnLeuProLeuTrpGlyMetThrIleGluGlnSerGluAsp	817
Db	2380	CAC TTC CGG AAT CCG GGG C TC TCT C C C T C C C A A G C A A G C T G T G A A A A G G A T A A C	2439
Qy	818	GluTrpGlyValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAla	837
Db	2440	GAGTGGTGTGTTCCACACTGTTTCAACATCTACCGCGGTCAAGAAACATGTGTGTGCA	2499
Qy	838	AlaSerSerArgSerGluMetGluLysTrpValGluAspIleGlnMetAlaIleAspLeu	857
Db	2500	GCCAGACATCGGCTGAGAAAGAAAGATGTGTGACCTGGAACCTCGCATCCAAGCA	2559
Qy	858	AlaGluLysSerSerSerProAlaProGluPhe-----LeuAlaSerSerPro	873
Db	2560	GCCAAAGATGGCGGTGACACGAGCCCTGTCACTGCCAGGCGGCACTGTGTGCTGCTCC	2619


```

Db      901 GGTAGCAGATGATGATGAAGACTTGTGAAGATATGTGTGAGTACCATACCTTTT    960
Qy      317 ArgLeuphegluGlProlyProlyProlyProValLeuPheSerArgLySerSer  336
Db      961 AGACTCTGTACACGCTTACGCAAGGCTGTCTTTCACGCGAGGCTCCCTCC    1020
Qy      337 PheArgPheSerGlyArgThrGlnLeuGlnValLeuAspTyrValLysGluGlyLysHis  356
Db      1021 TTCAGATACAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  1080
Qy      357 LysLysValGlnPheGluArgLysHisSerLysValHisSerLeuArgSerLeuAlaSer  376
Db      1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  1119
Qy      377 GlnProthrGluLeuAsnSerGluValLeuGluGlnSerGlnGlnSerThrSerLeuThr  396
Db      1120 -----ACATCTCTTCATGCTCTGACTGTAGTCTGCTTAACAGAGGCTCTTCAC  1173
Qy      397 PheGlyGluGlyAlaGlySerProGlyGlyGlnSerCysArgArgGlyLysGluProlys  416
Db      1174 GATGCTTGAGAACTTGCC-----TTC    1197
Qy      417 ValSerAlaGlyGluProGlySerHisProSerProAlaProArgArgSerProAlaGly  436
Db      1198 CTGCTTCAGCAATGTCTCTTATATCCACCCCTAGTCTTCTTGTCCCTCTGCT    1257
Qy      437 AsnLysGlnAlaAspGlyAlaAlaSerAlaProthrGluGluGlnGlnValLys  456
Db      1258 CTGCCCAATTGTAAGAGACAGACAGACAGCTCCTTGATCCGACAGGCTCCGCTCAAG  1317
Qy      457 Asp-----ArgThrGlnGln    461
Db      1318 AGCAGACAGACAGAGAGACAGCGACCATCATCATCTGACGCCCGACAGACAGCTCG  1377
Qy      462 SerLys---ProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro  478
Db      1378 GCCCATCTCCCTGGACCCCGGTGCTCGGCTGTCCAGGCTTTTCATGATAGTCT    1437
Qy      479 HisLeuSerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeu  498
Db      1438 CAGCTTCTCCCTCAGCTCAAGACCACTGACCTGTCCT-----GAGCTT    1488
Qy      499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuLysSerProLeuLysn  518
Db      1489 CAGGCCCGCTGAGTACAGCTGACGAGGGTGTATCCCTGTGCTCAGCCGTGTCTAGT  1548
Qy      519 AspGlnAlaCysProArgThrAspAspGluAspGlyArgGlyLysArgPheProThr  538
Db      1549 GGCCTGTGTACAGCAGATGACCAACAGAAAGACGAAACACAGANNNNNNNNNNNN  1608
Qy      539 AspLysAlaTyrPheHisAlaLysGlnValSerThrThrGluArgThrTyrLeuLysAsp  558
Db      1609 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  1668
Qy      559 LeuGluValIleThrSerThrPheGlnSerThrValSerLysGluAspAlaMetProGlu  578
Db      1669 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  1728
Qy      579 AlaLeuSerLeuLeuPheProAsnPheGluProLeuHisLysPheHisThrAsnPhe  598
Db      1729 GCCCTGATGCGCTGCTTTTCTCAACATGTATCAGTCTCAGAGTTCCACAGAGGCTTC  1788
Qy      599 LeuLysGluLysGluGlnArgLeuAlaLeuTyrGluGlyArgSerAsnAlaGlnLysArg  618
Db      1789 CTTCAACGAGTGTGAACAGAGCTGGCATCTGGGAAGGGCTCCAGTCCAGGCCACTTAATA  1848
Qy      619 ---AspTyrGlnArgGlyLysAspValMetLeuLysAsnIleGlnGlyMetLysHisLeu  637
Db      1849 GGTATATCCAGCAATCGGGAATCTCTCTCAGGAACATGCGCTTAAGGAATTT    1908
Qy      638 AlaAlaHisLeuLysPheHisSerGlnAlaLeuGlnAlaLeuGlnAlaSerGlyLysSer  657

```

```

Db      1909 ACTAGTACTTCCAAAGACAGATGATGCTCTTAACAGAACTGTGAAGGCCACAAACAC  1968
Qy      658 SerArgArgLeuGlnAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuPro  677
Db      1969 TGTAAAAAGCTGGAGGAGCTTACAAAGATTGAGCTCCAAAGGTTGTCTACTGCTCT    2028
Qy      678 LeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisTyrLysGlnValLeuGln  697
Db      2029 CTCAAACATCTCTGCTGAAGCCCGTCCAGAGGCTATGCTCAGTCCGCTGCTGTAGC    2088
Qy      698 ArgLeuCysLysHisHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeu  717
Db      2089 CGGCTGTGTCTACTCTCTCTGAGCACCGCACTATGCGCACTGCTGCAGAGGCACTG    2148
Qy      718 AlaGluIleThrGluMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPhe  737
Db      2149 AAGGCATCAGAGAGGACACCGAGCTCCAGCAAGGCTTACCGGCTGGAATACTTA    2208
Qy      738 GlnLysLeuHisGlnLeuLysLysAspLeuIleGlyIleAspAsnLeuValProGly  757
Db      2209 CAGAAATTGACGAGCTACAGCGAGACTGCTGCTGTGTAAGAAACCTATGCTCTGCTGG  2268
Qy      758 ArgGluPheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnArg  777
Db      2269 AGGAGATTATCCGTGAGGCTGCTGCACAAAGCTCACAAAGAAAGGCTCTCAACAGAG  2328
Qy      778 MetPhePheLeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrHisSerAsn  797
Db      2329 ATGTTTTCTGTCTCAGATATGTACTGTATACAGCAAAAGTGTACAGAGGCCAGT    2388
Qy      798 GlnPheLysValHisGlyGlnLeuProLeuTyrGlyMetThrIleGluGlnSerGluAsp  817
Db      2389 CATTTCCGATCCGTGCTTCTTCCACTCCGTGTGCTGCTGTGTAAGAAAGTGAAGAT  2448
Qy      818 GluTyrGlyValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleLeuValAla  837
Db      2449 GAATGCTGTGTTCTTCAATGCTTACCATGTACAGCTCAAGAAACATATGTGTAGCA  2508
Qy      838 AlaSerSerArgSerGluMetGluLysTyrValGluAspIleGlnMetAlaIleAspLeu  857
Db      2509 GCCAGCACTCGCTAAGAAAGAAAGTGTATGAGAGACCTGATGACGACATCAAA---  2565
Qy      858 AlaGluLysSerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLys  877
Db      2566 GCAGCAGAGTATCGGTGACTACCCCCAGTGTCTGTGGAGGCCGCTGTATATCTGT    2625
Qy      878 SerProAspGluAlaThrAlaAlaAspGlnLysSerGluAspLeuSerAlaSerArg  897
Db      2626 ACCCTGATCTTCTGATGAAGTCTCTGTGAAGAATCAGAAATGTCGAGAAACCGG  2685
Qy      898 ThrSerLeuGluArgGlnAlaProHisArgGlyAsnThrMetValHisValCysTyrHis  917
Db      2686 GGCCTCCGTGAGGGGAACAGCAGACCGGCTCAATCAACAAATGCTGTGTGTATC    2745
Qy      918 ArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerGly  937
Db      2746 CGTATATCAAGTGTGTCCAGACCAACAGTCAAGTGTGTGGNNNNNNNNNNNNNNNN  2805
Qy      938 AsnLeuLeuArgLysPheLysAsnSerAsnGlyTyrGlnLysLeuTyrValPheThr  957
Db      2806 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  2865
Qy      958 AsnPheCysLeuPhePheTyrLysSerHisGlnAspAsnHisPProLeuAlaSerLeuPro  977
Db      2866 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  2925
Qy      978 LeuLeuGlyTyrSerLeuThrIleProSerGluSerGluAsnIleGlnLysAspTyrVal  997
Db      2926 CTATTTGGTTTACACGCGAGGCTCTCCAGAGGCTACACAGATCCAAAGACTATGTC    2985
Qy      998 PheLysLeuHisPheLysSerHisValTyrTyrPheArgAlaGluSerGlyTyrThrPhe  1017
Db      2986 TTCAAGCTCCAAATTCAATCTCACTCTTCCGGGCTGAGACAAAGTACATATTT    3045

```

OY	1018	GLUARGTTPMTEGLUValIleArgSerIaThrSerSerAlaSerArgPro	1034
Db	3046	GAAGGTGAATGAGACGTATCAAAAGGCGACAGCTACCAAGGAGACC	3096
RESULT 9			
LOCUS	AY415516		
DEFINITION	AY415516	3126 bp	DNA
	Pan troglodytes PAR2 gene, VIRUAL TRANSCRIPT, partial sequence,		linear GSS 17-DEC-2003
ACCESSION	AY415516		
VERSION	AY415516.1		
KEYWORDS	GSS:		
SOURCE	Pan troglodytes (chimpanzee)		
ORGANISM	Pan troglodytes		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homidae; Pan.		
	1 (bases 1 to 3126)		
	Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,		
	Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,		
	Ferriera,S., Wang,G., Zheng,X.H., White,T.D., Shinsky,J.J.,		
	Adams,M.D. and Cargill,M.		
	Interfering nonneutral evolution from human-chimp-mouse orthologous		
	gene tios		
TITLE	Science 302 (5652), 1960-1963 (2003)		
JOURNAL	14671302		
PUBLISHED	2 (bases 1 to 3126)		
REFERENCE	Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,		
AUTHORS	Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,		
	Ferriera,S., Wang,G., Zheng,X.H., White,T.D., Shinsky,J.J.,		
	Adams,M.D. and Cargill,M.		
	Direct Submission		
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,		
	Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering		
	them based on alignment.		
FEATURES	Location/Qualifiers		
source	1..3126		
	/organism="Pan troglodytes"		
	/mol_type="Genomic DNA"		
	/db_xref="taxon:9598"		
gene	<1..>3126		
	/gene="PAR2"		
	/locus_tag="HCM583"		
ORIGIN			
Alignment Scores:			
Prod. No.:	9.27e-155	Length:	3126
Score:	1920.50	Matches:	424
Best Local Similarity:	53.0%	Conservative:	130
	40.5%	Mismatches:	439
Query Match:	35.2%	Indels:	53
DB:	14	Gaps:	13
US-09-555-342B-2 (1-1045) x AY415516 (1-3126)			
OY	11	GLYSerArgLeuGLYAlaLProGLYAsnSerGLYIleSerThrLeuGLYArgGLYGLnLys	30
Db	4	GGGACACCGCTTGAGNTGCCACAGCCNCTGTGAGNAGNACACCGCTTGAGCGGGAGCN	63
OY	31	ProProPothrProSerGLYLYsLeuValSerIleLysIleGlnMetLeuAspAspThr	50
Db	64	NNNNNGCCAGATGACAGAGAGAGCANTGTGCANNNTCAAGTAAAGCTGCTGACAACACC	123
OY	51	GlnGlnAlaPheGlnValLProGLNArgAlaProGLYLYsValLeuLeuAspAlaValCys	70
Db	124	ATGGAAATATTGTGACATTGACCTTAAATGCGATGGCCAGATTAATGACACAAGTGGG	183
OY	71	AsnHisLeuAsnLeuValGLnGLYLYsAspIyrPheGLYLeuGlnPheProAspHisLysLys	90
Db	184	AAGCGTTTAAACCTCGTAGAATGTGACTTTCGGGTGGAGTTTCAAAATATCTCATGCC	243

QY	91	llethraValTTPleAaBpLeuLeuLeuLeuProTleVallysgInlleArArpProLyshs	110
Db	244	TACGGAATTTGGCTTGAACCTATGAAACCCATTCATTAGCGAAATATCGAAGGCAAAAGAT	303
QY	111	ValValValLySpheValValLySphePheProPaaPPhIsthrGlaLeuGInglu	130
Db	304	GTGGTGTCTCCCTAGCTGTAAATTTTTCACCTGATCTCTGGTCACTACMAAAGA	363
QY	131	LeuthArGlyrLeuPheAlaLeuGInValLysgInaBpLeuLaGInGlyArLeuthr	150
Db	364	TATACAAATATCTGTGTGGCTTGCACACTTAAGAGAGACCTGTGAAGAGGCTTTGACC	423
QY	151	CysaBnaBpThrSerAlaAlaLeuLeuLeuSerHisAlaValGInSerGInlleLyap	170
Db	424	TGTGTGCACACCAACGCGGCCCTTCTCAATCCATCTCCGACGTGGAATAATAGANNNT	483
QY	171	PheaBpGluAlaLeuAaBpArGluHHisLeuAlaLySaBnLySTyrIleProGInGInaBp	190
Db	484	TACAGTGAACGCTGGACCGANNNNNNCTMAAAGGAAGCGAATTTGCTGGCCGACGCG	543
QY	191	AlaLeuGInaBpLySileValGluPheHisIAsnHisIleGlyGInThrProLaGlu	210
Db	544	NAATGCCCTGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGGCAACACCTGTGAG	603
QY	211	SerAaBpGInLeuLeuGInlleAlaArArGleuGInMetTyrGlyIleArGleuHis	230
Db	604	TCAGATTTCGAGGTGCTCGAATTTGCTCGAAATTTGGAAAGGTACGCGATCAGATTTCAC	663
QY	231	ProAlaLySaBpArGluGluTThrLySileAsnLeuAlaValAlaAsnThrGlyrIleLeu	250
Db	664	ATGGCTTCTGACAGAGAAAGAACCAAGATTCACTGGCAGATTTCACATGGGTGTACTC	723
QY	251	ValPheGInGlyPheThrLySileAsnAlaPheaBnTrpAlaLySaValArGlyLeuSer	270
Db	724	GTGTTCCAGNN	783
QY	271	PheLySaArGlySaArPheLeuIleLyLeuArGrProaBpAlaAsnSerAlaTyrGInaBp	290
Db	784	NNNGACCTTACAGAGAC	843
QY	291	ThrlaGluPheLeuMetAlaSerArArpPheCysLySerPheThrLySileGlyVal	310
Db	844	ACATTAGATATTTGTTGGTAGTAGATATATGTAAACCTTGGAAAGATTTGTGTG	903
QY	311	GluHisHisAlaPhePheArGleuPheGluPheGluPheLySProLySProLyPheValLeu	330
Db	904	GAGATATCACACTTTTTAGACTTTTGGACCAACTAAGCCAAAGCAAAAGCCGCTTTC	963
QY	331	PheSerArGlySerSerPheArGrPheSerGlyArGThrGInLySglInValLeuAaPtyr	350
Db	964	TTCAGCCGGGCTCCTCCTCAGATACAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1023
QY	351	ValLySglInGlyLyHisLySglValGInPheGluArGlySHisSerLySileHis---	369
Db	1024	NNNAAGCAGCAAGCAACCACATG	1083
QY	370	SerIleArGSerLeu--AlaSerGInProThrGluLeuAaSerGluValLeuGInGlu	388
Db	1084	TCCGTTCAGCTTCGATCGACGACACTTACCAAAACAG-----	1115
QY	389	SerGInGInSerThrSerLeuThrPheGlyGluGlyAlaGInSerProGlyLyGInSer	408
Db	1120	-----AGCATCTCATTCGCCAGGAGATTAAGAGACTCTGCGCTCCCATCT	1166
QY	409	CysArArG-----GlyLySgluPheValaSerAlaGlyGluProGly	423
Db	1165	TCAGCGATGCTTTTACTGCTCTCCCTCCACACTGTGTCCTGCGCTGGCTGGCAGAG	1224
QY	424	SerHisPProSerProAlaProArG-----ArGSerProAla	435
Db	1225	TTTAAAGCACACAGAGCTCCCGCAGATCCCAAGGTTTCTTCAATCAAAAGTCTCACT	1286
QY	436	GlySaBnLySglInAlaBpGlyAlaAlaSerAlaProThrGrInGInGluGInGluValaVal	455

```

Db      1285 GCAGAGAGACAGTGCAGTGCAGGCCCCGAC----- 1323
Qy      456 LysAspArgThrGlnGlnSerIleProGlnProP-----GlnProSerThrGly 472
Db      1324 -----ACACCATCGGCCAGCCCTCGGCGCCCCCGCATCCAGCTCGTCCAGGC 1374
Qy      473 SerLeuThrGlySerProHileuSerGluIleuSerValaanserGlnGlyVala1a 492
Db      1375 CTTTCCACGAGAGAGCTTCAGCTTCTCCCTCCAGCCGGAAGAGCCCCCTAGCTGAGC 1434
Qy      493 ProIlaenValThrLeuSerProAluIleuSerProAspThrLysGlnIlaSerProleu 512
Db      1435 CCTGCA-----TTTCAGGTGCTTTGGGCGCCAGCTGAACAGGGCTCATCCCACTC 1485
Qy      513 ILeSerProleuLeuAenAspGlnIlaCysProAlaGthrAspArgGluAspGluGlyArg 532
Db      1486 CTGAGCCCTGCTTCAGATGATGCTGGCGAGCCGGGATGACTACGAG--GAGCCCA 1542
Qy      533 ArgLysArgPheProThrAspLysAlaThrPheIlaIaLysGluValSerThrGlu 552
Db      1543 CACAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1602
Qy      553 ArgThrTyrlleuLysAspLysGluValIleThrSerTrpPheGlnSerThrValSerLys 572
Db      1603 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1662
Qy      573 GluAspAlaMetProGluIlaIleuLysSerIleuIlePheProAlaPheGluProleuHis 592
Db      1663 GAGACCGCATGCTGTGACTCTGATGACGCTCTTCCCAATCATGATCCCATCAT 1722
Qy      593 LysPheHleThrAenPheLeuLysGluIleGluGlnArgLysAlaIleuTrpGluGlyArg 612
Db      1723 GAGTTCCACAGAGGCTTCTCGCGAGGTGAGAGGAGCTGCACTTGGAAAAGGGCCC 1782
Qy      613 SerAenIlaGlnIleArg--AspTrpGlnArgIleGlyAspValMetLeuLysAsnIle 631
Db      1783 TCCAAAGCCCACAAAAGGAGCATCATCAAGAAATCGGGAGATCTGTCAGAAATCG 1842
Qy      632 GlnGlyMetLysVhIleuAlaIaIleuTrpLysVhIleuSerGluIlaIleuLys 651
Db      1843 CGCAGATTAAGGAGTTTACAGCTACTTCCAAAGACATGACGAGTCTTCAAGAAAT 1902
Qy      652 GluAsnGlyIleLysSerSerArgArgLysGluAsnPheCysAspArgPheGluGln 671
Db      1903 GAAAGGCTACCAACGCTGTAAAGAGTTGAGGAGCATGTCAAGAGAGTGAAGTGCAG 1962
Qy      672 LysValCysTyrlleuProleuAenThrPheLeuLeuArgProleuHisArgLeuMetHis 691
Db      1963 AAGGTGCTACTTGCCTCAACAGCTTCTGCGTGAAGCCCATATCCAGGCTGTCAGC 2022
Qy      692 TyrlleuGlnValLeuGluArgLysCysLysVhIleuTrpProSerHleValAspPheArg 711
Db      2023 TACGCTGCTGCGCGCGCTGTAGCGAGCATTAACCCCGGGCACATGACTACGCT 2082
Qy      712 AspCysArgIlaIaIleuAlaGluIleThrGluMetValaIaGlnIleuHisGlyThrMet 731
Db      2083 GACTGCATGACGCGCTGAAGCCATCAAGAGGTGACCAACCACTACAGACATTTCTC 2142
Qy      732 ILeuMetGluAenPheGlnLysLeuMetHisGluLeuLysAspLysIleGlyIleAsp 751
Db      2143 ATCCGCTGCGAAGACTGCAAGAGCTTAACGAGGCTGACGCGGAGCTGTGGGTAGAG 2202
Qy      752 AsnLeuValaIaProGlyArgGluPheIleArgLysSerLeuSerLysLeuSerGly 771
Db      2203 AACCTCTTGTCTCTGCGAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2262
Qy      772 LysGlyLeuGlnGlnArgMetPhePheLeuPheAsnAspValLeuLeuTyThrSerArg 791
Db      2263 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2322
Qy      792 GlyLeuThrAlaSerAenGlnPheLysValaIleGlyGlnLeuProleuTyrlleuMetThr 811

```

```

Db      2323 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2382
Qy      812 ILeGluIleuSerGluAspGluTrpGlyValProHileCysLeuThrLeuArgGlyGlnArg 831
Db      2383 GTGAAAGAAAGTGTAAACAGATGCTCTGTTCACACATGTTTACACATTTACCGGCTCAG 2442
Qy      832 GlnSerIleIleValaIaIaIaSerSerArgSerGluMetGluLysTrpValaGluAspIle 851
Db      2443 AAAACATCTGCTGTGGAGCCAGCACTCGCTGAGAAAGAAAGTGGATCTGTGACTG 2502
Qy      852 GlnMetAlaIleAspLysAlaGluLysSerSerSerProAlaProGluPhe----- 868
Db      2503 AACTCCGGATCCAAGAGCAAGATGCGCGGTGACANNCCCTTGCACTCCAGCGCGC 2562
Qy      869 ---LeuIlaSerSerProProAlaAspAluLysSerProAspGluAlaThrAlaIaAspGln 887
Db      2563 ACTGCTGCACTGCTCCCTCC-----AGATCCCCCAACGAG--GTAATCTGAGAGAG 2613
Qy      888 GluSerGluAspAspLysSerAlaSerArgThrSerLeuGluArgGlnAlaProHileArg 907
Db      2614 GAGTCNAGAGTATGCTCGCGGTGTCGCGAGCTCCCTGAGGGGAGCATGGCCAGACCGG 2673
Qy      908 GlyAenThrMetValaIaIaIaCysTrpHileArgAenThrSerValSerMetValaAspPhe 927
Db      2674 GCCAACACACCAAGATGACGCTGCTGATCCGGAACACACGCTGTCACAGGCGAGACAC 2733
Qy      928 SerIleAlaValaGluAenGlnLeuSerGlyAsnLeuLeuArgLysPheLysAsnSerAsn 947
Db      2734 AGTCACACTGTGAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2793
Qy      948 GlyTrpGlnLysLeuTrpValaIaPheThrAspPheCysLeuPhePheTyrlleuSerHis 967
Db      2794 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2853
Qy      968 GlnAspAsnHisProleuAlaSerLeuProleuGluTyrlleuSerThrIleProSer 987
Db      2854 NNNNATAGATCCACACTGCGCAGCTCCCGTGTGCGGTACANNNNNAGATCCCAAG 2913
Qy      988 GluSerGluAsnIleGlnLysAspTrpValaPheLysLeuHisPheLysSerHisValaTyr 1007
Db      2914 GAGGCCATGATGCAACAAAGACTATGTTTCAAGCTCCAGTCAATCCACAGTCTAC 2973
Qy      1008 TyrPheArgIlaGluSerGluTyThrPheGluArgTrpMetGluValaIleArgSerAla 1027
Db      2974 TTTCTCCGGGCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3033
Qy      1028 ThrSerSerAlaSerArg 1033
Db      3034 NNCAGTCAAGCGGGAGG 3051

RESULT 10
BX364606
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo
1 (bases 1 to 968)
Ll,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30370819.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

```


Db 329 CTCTACGGCATGACTATGAGAAAGTGAAGAGGATGGGGCGTCCCGCACTGCTGACT 388
 QY 827 LeuArgGlyValnArgGlnSerIleValAlaIleSerSerArgSerGlyMetGlyLeu 846
 Db 389 CTCGGGGGCGAGCGCATGCTCATGTGTGGCCGCGCATTTCTGGCCGGAGATGGACAG 448
 QY 847 TTPValGluAspIleGlnMetAlaIleAspLeuAlaGluLeuSerSerSerProAlaPro 866
 Db 449 TGGGTGAGACATACATGATGGCATTTGATGGCCGAGAAAGACAGCGGCCCGCCCC 508
 QY 867 GluPheLeuAlaSerSerProProAspAsnLeuSerSerProAspGlyAlaIleAlaAsp 886
 Db 509 GAGTTCCTGGCCAGACAGCCCGCGAGAACAGTCCCGCGAGAGGCG--GGGGGCGAC 565
 QY 887 GlnGluSerGlnAspAspLeuSerSerAlaSerArgThrSerLeuGlnArgGlnAlaProHis 906
 Db 566 CAGGAGTCGAGAGACACTCATGTGCTCGCGCACTTCGCTGAGAGCCGCGGCCGAC 625
 QY 907 ArgGlyAsnThrMetValHisValCysTrpHisArgAsnThrSerValSerMetValAsp 926
 Db 626 CGCGGCAACACATGATGCAAGTATGCTGGACCCGCAACACAGCGTGTCCATGTGTGAC 685
 QY 927 PheSerIleAlaValGluAsnGlnLeuSerGlyAsnLeuLeuArgGlyPheLeuAsnSer 946
 Db 686 TTCAGCGTGGCGGTGGAGAAATCAAGTGTCTGGGAACCTGTGAGGAATTCAAAAACAGC 745
 QY 947 AsnGlyTrpGlnLeuLeuTrpValValPheThrAsnPheCysLeuPhePheTrpLeuSer 966
 Db 746 AACGGGTGGAGAAAGCTGTGGGTGGTCTTACCAAGACTTTCCTCTTCTACAAATCG 805
 QY 967 HisGlnAspAsnHisProLeuAlaSerLeuProLeuGlnGlyTrpSerLeuThrIlePro 986
 Db 806 CACGAGACATACACCCCTTGGCAGCGTGTCTGTGCTGATGCTGCTACCATGCTCCC 865
 QY 987 SerGluSerGluAsnIleGlnLeuAspTrpValPheLeuLeuHisAspLeuSerHisVal 1006
 Db 866 TCGAGATCCAGAGCATCCACAAAGATTACGTGTTCAGCTGCATTTCCAG-TCCACAGTG 924
 QY 1007 TyrTrpPheArgAlaGluSerGluTrpThrPheGluIuArgTrp-MetGluValIle-ArgS 1026
 Db 925 TACTACTTCCGGGCGCGAAGCGAATACAGTTTGAAAGGTGGAATGAGTGATCCGAA 984
 QY 1026 eraIatnSerSerAlaSerArgProHis 1035
 Db 985 ATGCAACCACTCGGCTCACGTCCAC 1013
 RESULT 13
 BM551766
 LOCUS 1026 bp mRNA linear EST 20-FEB-2002
 DEFINITION AGENCOURT 6575294 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5479350
 5', mRNA sequence.
 ACCESSION BM551766
 VERSION BM551766.1 GI:18789071
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo
 1 (bases 1 to 1026)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strauberg, Ph.D.
 Email: cgarbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNCM2000 row: h column: 07

FEATURES High quality sequence stop: 675.
 Location/Qualifiers
 1..1026
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5479350"
 /issue_type="astrocytoma grade IV, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_11b="NIH MGC 98"
 /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCAAGAG (G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 7,348-123 Length: 1026
 Score: 1544.00 Matches: 318
 Percent Similarity: 93.0% Conservative: 2
 Best Local Similarity: 92.4% Mismatches: 13
 Query Match: 28.3% Indels: 11
 DB: 2 Gaps: 4
 US-09-555-342B-2 (1-1045) x BM551766 (1-1026)
 QY 667 AspPheGluLeuGlnLeuValCysTrpLeuProLeuAsnThrPheLeuLeuArgProLeu 686
 Db 3 GACTTGAGCTGACAGAAAGTGTGTACTTACCGCTCAACACTTCTCTCGCGCACTG 62
 QY 687 HisArgLeuMetHisTrpGlnValLeuGluArgLeuCysLeuHisHisProProSer 706
 Db 63 CACGGCTCATGACATCAACAGAGGTCTGGAGCGGCTGTGCAACACACCGCGAGC 122
 QY 707 HisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIleThrGlnMetValAlaGln 726
 Db 123 CAGCGGACTTCAAGGAGCTGCGAGCGCGCTTGGCAGAGATCACGAGATGCGTGGCAG 182
 QY 727 LeuHisGlyThrMetIleLeuMetGluAsnPheGlnLeuHisGluLeuLeuValAsp 746
 Db 183 CTCGAGGTATGATGATCAAGATGAGAAATTTCCAAAGCTGACCAACCTCAAGAAAGAT 242
 QY 747 LeuIleGlyTrpAspAsnLeuValValProGluArgGluPheTrpLeuGlnSerLeu 766
 Db 243 TTGATTGGCATTTGACATCTTGTGTTCGGGAAAGGAGTTCAATCCGCTGGCAGCCTC 302
 QY 767 SerLeuLeuSerGlyLeuGlnGlnArgMetPhePheLeuPheAsnAspValLeu 786
 Db 303 AGCAAGCTCTCGGGGAGAGGGGCTCCAGCAGCGCATTTCTTCTGTTCAACGACGTCCTG 362
 QY 787 LeuTrpThrSerArgGlyLeuThrAlaSerAsnGlnPheLeuValHisGlyGlnLeuPro 806
 Db 363 CTATACACGAGCGCGGGGCTGACGGCTCCATTCAGATTAAAGTTCACCGGCGAGCTCCG 422
 QY 807 LeuTrpGlyMetThrIleGluGluSerGluAspGluTrpGlyValAlaProHisCysLeuThr 826
 Db 423 CTTTATGGCATGACATGAGAGACGAAAGAGATGGGGGGGTGCCCATCTGAC 482
 QY 827 LeuArgGlyValnArgGlnSerIleValAlaIleSerSerArgSerGlnMetGlyLeu 846
 Db 483 CTCGGGGGCGAGGAGTCCATCATCTGTGGCCGCGCATTTCTGTGCTCCAGATGGAGAG 542
 QY 847 TTPValGluAspIleGlnMetAlaIleAspLeuAlaGluLeuSerSerSerProAlaPro 866
 Db 543 TGGGTGAGACATCAATGATGCGCATTTGACCTGGCGAGAGAGCAGCGGCCCGCTC 602
 QY 867 GluPheLeuAlaSerSerProProAspAsnLeuSerSerProAspGlyAlaIleAlaAsp 886
 Db 603 GAGTTCCTGGCCAGAGCGCCCTGACAAAGTCTCCGTGATGAGCCACCGGCGCTGAC 662

Oy	887	Glnc1usSerg1uAaPAPLeuSer1AaSerAaGTHrSerLeuGlu1uAIGl1A1AProHs	906
Db	663	CAGGAGTGAAGAGGAGACTTGAAGCCCTCGCGACATGCTGGAGCGCCAGGCCCGGAC	722
Oy	907	ArgG1AaEnThmEValH1eVal1CyETrPHs1AaGAnThrSe1ValSerMetValAsp	926
Db	723	CGCGGCAACACAAATGCTGACCTGTGCTGGCACCGCCAACACCAAGCTCTCCATGTTGAC	782
Oy	927	PheSer11eAlaVal1GluAaNG1uLeuSerg1yAaEnLeuAaUaLyPheLyAaAnSer	946
Db	783	TYCAGCATCGCAGTGGAGAAATCAGTGTCTGGAAACCTGTGAGAAATTCAAAAACAGC	842
Oy	947	Aen-G1YTrpG1uLyLeuTrp-Val1ValPheThrAaEn---PheCyLeuPhePheTyr	964
Db	843	AAACGGGGGGGACGAAGAGCTGTGGGGGTGGGTTCACAAAATTCTGGCTTTCTTCTAC	902
Oy	965	LySserH1aGln---AspAaHs1AProLeu1AaSerLeuProLeu-LeuG1YTrSerLe	983
Db	903	AAATACACCAAGGACATTACCTCCCTTCCCGACCTGCTGCTGGCTTA---CT	959
Oy	983	uThr1LeProSerg1uSerg1u-----AaH1eG1uLyAaP1YrVal1-PheLyL	1000
Db	960	CGGTTCAACATCCCTCCCTCGAAGTCCCAAGACCTTCAGAAAGACTAGTGTTCAGAC	1019
Oy	1000	eu	1000
Db	1020	TT	1021
RESULT 14			
LOCUS	BO881649	912 bp	mRNA linear EST 16-AUG-2002
DEFINITION	AGENCOURT_8728249 NIH_MGC_47 Homo sapiens CDNA clone IMAGE:6339275		
ACCESSION	BO881649		
VERSION	BO881649.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens (human)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 912)		
AUTHORS	NIH-MGC http://mgs.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Stransberg, Ph.D. Email: csapbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://lmgs.llnl.gov Plate: LLCM2533 row: n column: 12 High quality sequence stop: 684. Location/Qualifiers 1..912 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone_image="6339275" /tissue_type="neuroblastoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G) size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit		

ORIGIN		(Stratagene) and Superscript II RT (Life Technologies).	
		Note: this is a NIH_MGC Library."	
Alignment Scores:			
Pred. No.:	2,44e-120	Length:	912
Score:	1514.00	Matches:	296
Percent Similarity:	97.7%	Conservative:	1
Best Local Similarity:	97.4%	Mismatches:	7
Query Match:	27.7%	Indels:	1
DB:	3	Gaps:	0
US-09-555-342B-2 (1-1045) x BQ881649 (1-912)			
QY	182	LYSAENLYSYTYLLEPROGLINGLNSAPALALEUGLNASPYLEIILEVALGLUPHENISHLS	201
DB	1	AAAAATAATATCATATCCTCAGCAAGCCGACCTTAGAGGACAAATCTGGATTTCAACAT	60
QY	202	ASNHISIIIEGLYGLNTHPROLAGLUSERASRPHEGLNLEULEUGLIULEALATGAG	221
DB	61	AACCAATGTCACAAACACACAGCAATATCAGATTTCCACTCTTAAGATTCGCCCTCGG	120
QY	222	LEUGIMECTYRGYLYLEARGLEUHSIPROALYASAPARGGLNGLYTHRYSLLEAS	241
DB	121	CTAGAGATGATAGGAATCCGATGTGACCCGGCCAGAGACAGGAGAAAGCCAGGAATCAAT	180
QY	242	LEUALAVALAIAASTHCRYLILEUVALIPHEGLNGLYPHETHRTRYSLLEAENALAPHE	261
DB	181	CTGGCCGTCGCCACACCGGAATTTCTAGTGTTCAGAGGTTTCACTAAGATCAATGCTTC	240
QY	262	ASNTTPALALYASVALARGLYLEUSERPHELYSARGLYSARGPHELEULIELYLSLEARG	281
DB	241	AAC TGCGCCAGAGTCCGAGACCTGAGCTTCAAGAGGAACGCTTCTCATCAAGCTCCGG	300
QY	282	PROASAPALAAENSERIATRYCTINASPTTLEUGLNPHELEUMERLASERTARGAPHE	301
DB	301	CCAGATGCCAATAGGCGGTACCAAGATACCTTGGAATTCCTAATGCGCAGTCGGATTTTC	360
QY	302	CYSLYSERPHEPTRYLYIECYVALGIUHSIIISALAPHEPHEARGHEUPHEGLINU	321
DB	361	TGCANGTCTTCTGAAATATCTGTGTGAACATCATGCTTCTTTTGACTTTTGGAAAG	420
QY	322	PROLYSPROLYSPROLYSPROVALLEUPHESEARGLYSESERPHEARGPHESEGLY	341
DB	421	CCCAAACCAAGCCCAAGCCGCTCTTTTACCCGGGGGTCAATCATTTGGTTCAGTGGT	480
QY	342	ARGTHRGILNLYEGLNVALLEUASPTRYVALYEGLYGLYVHSIIISLYELYSVALGINPHE	361
DB	481	CGGATCTCAAGACGAGTCTCGACATATGTTAAAGAGGACATTAAGAGGTGCGATTT	540
QY	362	GLUATGLYVHSIIISERYLYIEHSERILEARGSERLEUALASERGLNPROTHRGILNLEU	381
DB	541	GAAAGGAAAGCACAGCAAGATTCATTTCCGGAAGCTTGTTCACAGCTTACAGAACTG	600
QY	382	ASNSERGIUALLEUGLUGLNSERGLNGLNSERTHSERLEUTHPHEGLYGLIUALA	401
DB	601	AATTCGGAAGTCTGAGACAGTCTCAGACAGACACACAGCTTACATTGAGAGAGGTGCC	660
QY	402	GLUSERPROGLYGLYGLNLSERCYSARGARGLYLYEGLNPROLYVALSERIAGLYGLU	421
DB	661	GAATCTCCAGGGGGGCGACAGCTGCGCGGAGGAAAGAAACCAAGGTTTCCGCCGGGGAG	720
QY	422	PROGLYSERHISPROSERPROALAPROARGATGSRERPROALAGLYASNLVSGIUALAAB	441
DB	721	CCGGGGTTCACCCGAGCTTGCGCCAGAGAAAGCCCGCGGGTAAACAAGCGCGGAC	780
QY	442	GLYLAALASERIALAPROTHRGILUGLUGLUGIUALVALIYASAPARGTHRGINGIN	461
DB	781	GGAAGCCGCTTCGGGGCCACCGAGGAAAAAGAGAGGTCTTTAAGATAGGACCCGACAG	840
QY	462	SELYSPROGLINPROPROGLINPROSERTHRGYLSERYLEUTHRGYLYSEPROHISLEUSER	481
DB	841	AGTAACTCAGGCCCGCAGCAAGACAGGCTCTCTGACATGGGAGTCT-CACCTTTTC	899

Qy 482 GluLeuSerVal 485
 |||||
 Db 900 GAGCTGTCTGTG 911

RESULT 15
 BU931960 904 bp mRNA linear EST 18-OCT-2002
 LOCUS BU931960
 DEFINITION AGENCOURT 10485886 NIH_MGC_109 Homo sapiens cDNA clone
 IMAGE:6672254 5', mRNA sequence.
 ACCESSION BU931960
 VERSION BU931960.1 GI:24120779
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 904)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LCM2952 row: h column: 14
 High quality sequence stop: 721.
 Location/Qualifiers
 1..904
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6672254"
 /tissue_type="teratocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_109"
 /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
 XhoI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Alignment Scores: 8.77e-119 Length: 904
 Pred. No.: 1496.00 Matches: 296
 Score: 98.3% Conservative: 1
 Best Local Similarity: 98.0% Mismatches: 2
 Query Match: 27.4% Indels: 3
 DB: 3 Gaps: 0

US-09-555-342B-2 (1-1045) x BU931960 (1-904)

Qy 323 LysProLysProLysProValLeuPheSerArgLysSerPheArgPheSerGlyArg 342
 |||||
 Db 1 AAACCAAGCCCAAGCCCGTCTCTTAGCCGGGGTCATCATTTCCGTTCAAGTGTCCG 60

Qy 343 ThrGlnLysGlnValLeuAspTyrValLysGlnGlyGlyHisLysLysValGlnPheGlu 362
 |||||
 Db 61 ACTCAGAGCAGGTTCTCGACTATGTAAAGAGAGACATAAGAAAGTGCAGTTGAA 120

Qy 363 ArgLysHisSerLysLysHisSerLysArgSerLeuLysSerGlnProThrGlnLeuAsn 382
 |||||
 Db 121 AGGAAGACACGACAGATTTCATTATCCGAGCCTTCTTCACAGCCTACAGAACTGAAT 180

Qy 383 SerGluValLeuGlnGlnSerGlnGlnSerThrSerLeuThrPheGlyGlyGlyValGly 402
 |||||
 Db 181 TCGAAGTGTGTGAGCAGTCTCAGCAGACACAGCCTTACATTTGAGAAAGGTGCCGA 240

Qy 403 SerProGlyGlyGlnSerCysArgArgGlyLysGlnProLysValSerLysGlyPro 422
 |||||
 Db 241 TCTCCAGGGGGCCAGAGCTGCCCGCAGAGAAAGAAACCGAAGGTTTCCGCCGGGAGCCG 300

Qy 423 GlySerHisProSerProLysProLysProLysProLysProLysProLysProLys 442
 |||||
 Db 301 GGGTCCGACCCGAGCCCTGCGCAGAGAGAGAGCCCGGGGTAAACAGCGGAGCGGA 360

Qy 443 AlaAlaSerAlaProThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 462
 |||||
 Db 361 GCCGCTCCGGCCGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420

Qy 463 LysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeuSerGlu 482
 |||||
 Db 421 AAACCTCAGCCCCCGCAGCCACAGCAGAGCTCCCTGACTGGCAGTCTCACTTTCCGAG 480

Qy 483 LeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerProAsnLeu 502
 |||||
 Db 481 CTGTCTGTGAATCCGAGGGGGAGAGTGGCCCTGCCAAGCTGACTTGTCTCCCACTG 540

Qy 503 SerProAspThrLysGlnAlaSerProLeuLysSerProLeuLysAsnArgGlnAlaCys 522
 |||||
 Db 541 AGCCCGACACCAAGCAGGCTCTCCCTTGATCAGCCGCTGTAATGACACGAGCCTGC 600

Qy 523 ProArgThrAspArgGlnAspGlyLysArgGlyLysArgPheProThrAspLysAlaTyr 542
 |||||
 Db 601 CCCCGACGAGCATGAGATGAGGGCCGAGAGAAAGATTCCCACTGATTAAGCGTAC 660

Qy 543 PheIleAlaLysGlnValSerThrThrGluArgThrTyrLeuLysAspLeuGluValIle 562
 |||||
 Db 661 TTCTATGCTTAAGAAAGTGTCTACACCGACGAACTATCTGAAGATCTCGAAGTTATC 720

Qy 563 ThrSerTyrPheGlnSerThrValSerLysGluAspAlaIleProGluAlaLeuLysSer 582
 |||||
 Db 721 ACTTCGAGGTTTCAGAGCAGACAGTGAACAAAGAGAGCGCATGCCGGAAGCAGTGAAGAT 780

Qy 583 LeuIlePheProAsnPheGlnProLeuHisLysPheHisThrAsnPheLeuLysGluIle 602
 |||||
 Db 781 CTCATATTTCCGAAATTTTGAACCTTGCACAAATTTTCATTAATTTCTCAAGGAATTT 840

Qy 603 GluGlyArg-LeuAlaLeu-TyrGlyGly-ArgSerAsnAlaGlnIleArgAspTyrGln 621
 |||||
 Db 841 GAGCAACGACCTCCCTGTGGGAAGGCCCGCTCAAAATGCCAAATCAGAGATTACAA 900

Search completed: May 20, 2006, 08:51:24
 Job time : 11281 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: May 20, 2006, 02:18:42 ; Search time 449 Seconds

(without alignments)
6532.212 Million cell updates/sec

Title: US-09-555-342B-2

Perfect score: 5463

Sequence: 1 MGIEQRPTGSRGLGAPENS.....SATSSASRPHVLSKESLVY 1045

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=x1h

-O=/abs/ABSSMB.spool/US09555342/runat_19052006_121500_13276/app_query.fasta-1

-DB=Issued_Patents_NA -OPMT=fastcap -SUFFIX=p2n.rn1 -MIMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt

-LIST=45 -DOCALL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTPR=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-HOST=abssmb2h -USBR=US09555342 -@CGN_1_1_204 -@runat_19052006_121500_13276

-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_NA:

1: /EMC_Celerra_SIDS3/prodata/2/ina/1.COMB.seq:

2: /EMC_Celerra_SIDS3/prodata/2/ina/5.COMB.seq:

3: /EMC_Celerra_SIDS3/prodata/2/ina/6.COMB.seq:

4: /EMC_Celerra_SIDS3/prodata/2/ina/6.COMB.seq:

5: /EMC_Celerra_SIDS3/prodata/2/ina/7.COMB.seq:

6: /EMC_Celerra_SIDS3/prodata/2/ina/7.COMB.seq:

7: /EMC_Celerra_SIDS3/prodata/2/ina/8.COMB.seq:

8: /EMC_Celerra_SIDS3/prodata/2/ina/8.COMB.seq:

9: /EMC_Celerra_SIDS3/prodata/2/ina/8.COMB.seq:

10: /EMC_Celerra_SIDS3/prodata/2/ina/8.COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	724	13.3	4292	US-09-949-016-1488 Sequence 1488, App
2	724	13.3	4336	US-09-949-016-208 Sequence 208, App
3	713	13.1	2853	US-09-949-016-1933 Sequence 1933, App
4	712.5	13.0	6263	US-09-664-958-5 Sequence 5, Appl
5	712.5	13.0	6268	US-09-566-921-57 Sequence 57, Appl
6	626	11.5	3398	US-09-799-451-201 Sequence 201, Appl
7	563	10.3	2872	US-09-906-779-3 Sequence 3, Appl
8	557.5	10.2	2493	US-10-104-047-41 Sequence 41, Appl

9	541	9.9	3984	US-09-848-294-1 Sequence 1, Appl
10	508	9.3	2156	US-09-949-016-3593 Sequence 3593, App
11	503.5	9.2	576	US-09-270-767-1995 Sequence 1995, App
12	503.5	9.2	576	US-09-270-767-17277 Sequence 17277, App
13	497.5	9.1	2931	US-10-094-749-1401 Sequence 1401, App
14	443	8.1	4272	US-09-949-016-5747 Sequence 5747, App
15	424.5	7.8	4092	US-09-566-921-115 Sequence 115, App
16	405.5	7.4	2170	US-09-854-133-728 Sequence 728, App
17	403.5	7.4	3438	US-10-164-595-29 Sequence 29, Appl
18	399.5	7.3	1979	US-09-949-016-4612 Sequence 4612, App
19	396.5	7.3	4080	US-08-446-345-35 Sequence 35, Appl
20	396.5	7.3	4081	US-09-949-016-1292 Sequence 1292, App
21	382	7.0	3044	US-09-814-915A-84 Sequence 84, Appl
22	382	7.0	3044	US-09-880-107-3718 Sequence 3718, App
23	359.5	6.6	2080	US-08-179-738-9 Sequence 9, Appl
24	359.5	6.6	2080	US-08-628-145-9 Sequence 15, Appl
25	359.5	6.6	2257	US-08-171-118-15 Sequence 15, Appl
26	359.5	6.6	2257	US-08-478-087-15 Sequence 15, Appl
27	343	6.3	1862	US-08-179-738-4 Sequence 4, Appl
28	343	6.3	1862	US-08-628-145-4 Sequence 4, Appl
29	338.5	6.2	1817	US-08-179-738-1 Sequence 1, Appl
30	338.5	6.2	1817	US-08-628-145-1 Sequence 1, Appl
31	338.5	6.2	1833	US-08-179-738-6 Sequence 6, Appl
32	338.5	6.2	1833	US-08-628-145-6 Sequence 6, Appl
33	337	6.2	310	US-09-621-976-8193 Sequence 8193, App
34	332.5	6.1	2723	US-10-104-047-1219 Sequence 1219, App
35	331	6.1	573	US-09-270-767-1542 Sequence 155, App
36	331	6.1	573	US-09-270-767-15447 Sequence 15447, App
37	322	5.9	56516	US-08-996-306-1 Sequence 1, Appl
38	322	5.9	56516	US-09-338-907-1 Sequence 1, Appl
39	322	5.9	56516	US-09-218-207-1 Sequence 1, Appl
40	322	5.9	56520	US-09-338-907-179 Sequence 179, App
41	322	5.9	56520	US-09-218-207-179 Sequence 1033, App
42	321	5.9	1626	US-09-620-312D-1033 Sequence 26, App
43	321	5.9	1699	US-09-913-171A-26 Sequence 1542, App
44	317.5	5.8	449	US-09-270-767-1542 Sequence 1542, App
45	317.5	5.8	449	US-09-270-767-16824 Sequence 16824, App

ALIGNMENTS

RESULT 1
US-09-949-016-1488
Sequence 1488, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1488
LENGTH: 4292
TYPE: DNA
ORGANISM: Human
US-09-949-016-1488

Alignment Scores:
Pred. No.: 4.58e-58
Score: 724.00
Percent Similarity: 44.0%
Best Local Similarity: 28.8%
Query Match: 13.3%
Length: 4292
Matches: 221
Conservative: 117
Mismatch: 287
Indels: 143
Gaps: 24

Qy 664 eCyArGAbpRhegIuLeuGInLySvAlCyEtYrLeuProLeuAenThrPheLeuLeuA 684
Db 2698 -----CCATTGTCGAACCTGAACCAAA 2720
Qy 684 gPro-----LeuH1aYrLeuMeH1eTyYrLySvAlValLeuGInLyA 700
Db 2721 ACCATACATATGATGTCGACAGATTGATGCGGGCGCTGGTGTGATTGCGGACCTTA 2780
Qy 700 gLySvH1eH1eProProSerH1s 707
Db 2781 CTGACCGCAAAACCATCATCAT 2802

RESULT 2
US-09-949-016-208
/ Sequence 208, Application US/09949016
/ Patent No. 681239
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C1001307
/ CURRENT APPLICATION NUMBER: US/09/949, 016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 208
/ LENGTH: 4336
/ TYPE: DNA
/ ORGANISM: Human
/ US-09-949-016-208

Alignment Scores:
Pred. No.: 4,67e-58 Length: 4336
Score: 724.00 Matches: 221
Percent Similarity: 44.0% Conservative: 117
Best Local Similarity: 28.8% Mismatches: 287
Query Match: 13.3% Indels: 143
Gaps: 24

US-09-555-342b-2 (1-1045) x US-09-949-016-208 (1-4336)

Qy 34 ThrProSerGlyLySvLeuValSerIleYrIleGInMeLeuAerPheRnGInGluA 53
Db 678 ACCAAGAAACCAAACTGTCAGTGAAGTGCCTTTAGTGGCAACCAATACAGC 737
Qy 54 PheGInValProGInATGATGATProGlyLySvAlValLeuAerPheRnGInGluA 73
Db 738 TGTGACCTGGAGAAACATGCCAAGGACCAAGTGTATTGACAAAGTGTGAACACATC 797
Qy 74 AsnLeuValGInGlyAerPheRnGInGlyLeuGInPheRnAerPheRnGInGluA 93
Db 798 AATCTCTTGGAGAAAGCTATTGACTTTTGTGAGAAAGCCTTGAGCAAGAAAC 857
Qy 94 TrpLeuAerPheLeuLySvProIleValIleGInIleAerPheRnGInGluAVal 113
Db 858 TGGTATGATCTCTGTAAGAAATTAAGACACCTGAAACCTTCATGGCTATTGACT 917
Qy 114 LySvPheValIleYrPheRnProProAerPheRnGInGInGInGInGluLeuThrA 133
Db 918 ---TTTATGTGAAGTTTATCTCTGATCTCTCAATTGATGATGATGATGATGATG 974
Qy 134 TyLeuPheAlaLeuGInValIleYrGInAerPheRnGInGluAerPheRnGInGluA 153
Db 975 TACTCTTGTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1034
Qy 154 ThrSerAlaIleLeuIleSerH1eIleValGInSerGInIleGlyAerPheRnGlu 173

Db 1035 GTGACCTATGCTCTCCCTGGGATCTCTACACCTGTGAGGCTBACTGTGTGACTATGAC--- 1091
Qy 174 AlaLeuAerPheRnGInGluH1s-----LeuAlaYrAerPheRnGInGluPro---GIn 188
Db 1092 -----CCAGAAAGCAATGGCAGCATGCACTGATGATTAATTCAGTTGGCCCTACTAC 1145
Qy 189 GInAerPheRnGInGluAerPheRnGInGluValIleGInPheH1eH1eAerPheRnGInGlu 208
Db 1146 ACTAAGAGCTGGAGAGAGAGGTGGCAGCTGCAAAACCCACAGGGGCTTATGCGCA 1205
Qy 209 AlaGInSerAerPheGInLeuLeuGInIleAlaYrGInGluMeTyYrGlyIleA 228
Db 1206 GCACAGACCTATTCCTCCGATCTTGAATAATGCAAAAGGCTTCCATGATGCTGAC 1265
Qy 229 LeuH1eProAlaYrAerPheRnGInGluIleYrIleAerPheRnGInGluAlaAerPheRnGly 248
Db 1266 CTACATCATGCCAAGGCTCAGAAAGTGTGACACTCAAGCTGGGCGTGTGTGCTAATGCA 1325
Qy 249 IleLeuValPheGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 268
Db 1326 CTCTCATTTTACAAAGACAGACTGCGATCATGTTTCTGTGGCCGAAATCTTAAA 1385
Qy 269 LeuSerPheLySvArGlySvArGlyPheLeuIleYrLeuAerPheRnAerPheRnAerPheRn 288
Db 1386 ATTTCTATTAAGCAGCTACTCTTCAATTAAGTCAAGCCGAGAGCTGGAACGTTT 1445
Qy 289 GInAerPheRnGInGluPheLeuMeTyYrAerPheRnGInGluAerPheRnGInGluAerPheRn 308
Db 1446 GAGAGTACCTATGATGATCAATGCCAACCCAGCCGAGCAAGAAAGTGTGAAAGT 1505
Qy 309 CyValGInH1eH1eAlaPheRnAerPheRnGInGluGInGluProLySvProLySvPro 328
Db 1506 TGGTGAGCATCATCTTCTACAGGCTGTT---TCTCAAGACGACCAACAAAGCC 1562
Qy 329 ValLeuPheSerAerGlySerSerPheAerPheSerGlyAerGInGInGInGInGInGInGIn 348
Db 1563 AAGTCTGACCTTGGGCTCAATTCGCTAATGTCGCGCCGACCCAGACAGACCCGC 1622
Qy 349 AsPTrValIleGInGluGlyIleH1sLySvAlGInPheGInGluAerPheRnGInGluAerPheRn 368
Db 1623 CAGGCGACGACCTCATGATGATGACGACGACACACTTGAAGGCACTTGTATGAACGG 1682
Qy 369 H1eSerIleAerPheRnGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 388
Db 1683 GTCTCC---AGAGTCTAATGAGCTCGATT-----GGTGTCAAGACCAA 1727
Qy 389 SerGInGInSerThrSerLeuThrPheGInGluGlyAlaGInSerProGlyGInSer 408
Db 1728 AGTCTTATGAG-----GATTTCTGGCGCTGCTGGG 1760
Qy 409 CyArGArGlyLySvGluProLySvAlSerAlGlyGluProGlySerH1sProSerPro 428
Db 1761 -----GAGATTTCAGCTTATGACCTGACT---GTGAGCAT 1796
Qy 429 AlaProAerGserProAlaGlyAerPheRnGInGInGInGInGInGInGInGInGInGInGInGIn 447
Db 1797 GCCGTGTACAAATGAGGCGGACAGGGAAGTGAAGCCCAATGAAGCCCAAT 1856
Qy 447 ----- 447
Db 1857 TTGCAGCTATTGAAGAAAGAAATTCCTTGAAGTGAAGGAGTAAATTATTATGTC 1916
Qy 448 -----ThrGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 457
Db 1917 AGACATAGCAATTTAATGTTGAGAACTGATTAAGCCAGAGAGCATATGTAAGAAAT 1976
Qy 458 ArgThr-----GInGInSerLySvProGInPhePro--- 467
Db 1977 CAGGCTAGCATTAAGTAACTCAAGCCAAATTTATGGAATCCACACTTGAAGCCGCGCCT 2036
Qy 468 -----GInPheSerThrGlySerLeuThrGlySerProH1s 479

Db 2037 AATGAATGGGAAAAACGACCTATACACCTCTGCTCCCTACAGACACAGGAGTTCACAT 2096
Qy 480 -----LeuSerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsn 495
Db 2097 GAGACTCTGAATATATGCTGGAGAGAGAGACCGGGGAGAGGTGGGAAAAGCGAAAGACTA 2156
Qy 496 ValThr-----LeuSerProAsnLeuSerPro--AspThrLys 507
Db 2157 ATCACAGAAAGAAATGAATGTGTAAGATATACCTCGAGGTGCTCTGGGAGATTCGT 2216
Qy 508 GlnAlaSerProLeuLe-----SerProLeuLeuAsnAspGlnAlaCyPro 523
Db 2217 AAGGTGAGGCTGTGGACACAAAAGACTCCCTCTGCTTCTTGAGACACACAGCC 2276
Qy 524 ArgThrAspAspGluAspGluGlyArgArgLysArgPheProThrAspLysAlaTyPhe 543
Db 2277 AGCAGTCAGAGTGAAG 2327
Qy 544 IleAlaLysGluValSerThrThrGluArgThrTyLeuLysAspLeuGluAlaIleThr 563
Db 2328 GTGACCGAGGGACCATCAGAGAGAGAGAGAGAGAGATGATAGAGAGGTGAGAGAAACC 2387
Qy 564 SerTrpPheGlnSerThrValSerLysGluAspAlaMetProGluAla----- 579
Db 2388 CGCCCGGACCCAGAGGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2447
Qy 580 -----LeuLysSerLeuLe--PheProAs 587
Db 2448 CAAGCAGGTCCAGTGAATCAGATGAAACAGTGTATCCAGAAAATGATAGTGCCCA 2507
Qy 587 nPheGluProLeuHisLysPheHisThrAsnPheLeuLysGluIleGlnGlnArgLeuAl 607
Db 2508 AAGATACCCGAG 2552
Qy 607 AlaLeuTrpGluGlyArgSerAsnAlaGlnIleArgAspTyr--GlnArgIleGlyAspVa 626
Db 2553 GACATGGGAG 2612
Qy 626 IMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaAlaHisLeuTrpLysHisSerGI 646
Db 2613 GACATGATGATTTTGGCCAAATATATTTGTTTTCAGAGCCACAGTGGTAAACAGAG 2672
Qy 646 ValAlaGluAlaLeu-----GlnAsnGlyIleLysSerSerArgArgLeuGlnAsnPh 664
Db 2673 ATGTAACAATTTCTGATGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2727
Qy 664 eCyAspArgPheGluLeuGlnLysValCyStrLysProLeuAsnThrPheLeuLeuArg 684
Db 2728 -----CAATGTGCCAAACTGAGCCAA 2750
Qy 684 gPro-----LeuHisArgLeuMetHisStrLysGlnValLeuGlnLysArgLeuCy 700
Db 2751 ACCATCAACATATAGTCTCCACAGATTTGATGGCGGGCTGTGTATTTGGGACAGTTA 2810
Qy 700 sLysHisHisProProSerHis 707
Db 2811 CTGACCGCACAAACCATCACAT 2832

RESULT 3
US-09-949-016-1933
; Sequence 1933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1933
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1933

Alignment Scores:
Pred. No.: 2,52e-57 Length: 2853
Score: 713.00 Matches: 190
Percent Similarity: 46.1% Conservative: 90
Best Local Similarity: 31.3% Mismatches: 197
Query Match: 13.1% Indels: 130
DB: 3 Gaps: 17

US-09-555-342B-2 (1-1045) x US-09-949-016-1933 (1-2853)

Qy 16 AlaProGluAsnSerGlyLeSerThrLeuGluArgGlyGlnLysProProThrPro 35
Db 719 AGCCCTCAATCAAAACAG 772
Qy 36 SerGlyLysLeuValSerIleLysIleGlnMetLeuAspAspThrGlnLysAlaPheGlu 55
Db 773 AAACAGAGAACATGACATGCAAGGTTCTTGTGGATGACACAGTTATATGAAATGTGTT 832
Qy 56 ValProGlnArgAlaProGlyLysValLeuLeuAspAlaValCyAsnHisLeuAsnLeu 75
Db 833 GTGGAGAAACATGCTAAG 892
Qy 76 ValGluGlyAspTyrPheGlyLeuGluPheProAspHisLysLysIleThrValTrpLeu 95
Db 893 TTGGAGAGAGAGATTTTGTGTGATGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 952
Qy 96 AspLeuLeuLysProIleValLysGlnIleArgArg--ProLysHisValValValLys 114
Db 953 GATTCGCCCAAGAAATAAAG 1006
Qy 115 PheValValLysPhePheProProAspHisThrGlnLeuGlnGluLeuThrArgTyr 134
Db 1007 TTAAATGTAAGATTATATACCTGACCCAGACAGTTAAACAGAGAGAGAGAGAGAGATAT 1066
Qy 135 LeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCyAsnAspThr 154
Db 1067 TATTTATGCTTTCAGTTCGGACAGACATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1126
Qy 155 SerAlaAlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAspGluAla 174
Db 1127 ACCTTAGCATTTATTAAGTTCTTACACCATCCAGTCTGAACCTGGAGAGACTACGACGA 1186
Qy 175 LeuAspArg--GlnHisLeuAlaLysAsnLysValTrpIleProGlnGln--AspAlaLeu 192
Db 1187 CTCATGCGGTGATATGTAATGTAATTTAACTGGCCCAATGAGCCAAAGAGACTT 1246
Qy 193 GlnAspLysIleValGlnPheHisHisAsnHisIleGlyGlnThrProAlaGlnSerAsp 212
Db 1247 GAAAGAGAGAGATGAG 1306
Qy 213 PheGlnLeuLeuGlnIleAlaArgArgLeuGluMetTyGlyTyLeuArgLeuHisProAla 232
Db 1307 TTGGAGTTCTTGAGAGATGCCAAAAGAGTTGATATGATGAGAGTTGATCTTCAATAAGCA 1366
Qy 233 LysAspArgGluGlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPhe 252
Db 1367 AAGGACTTGAAG 1426
Qy 253 GlnGlyPheThrLysIleAsnAlaPheAsnTrpAlaLysValArgLysLeuSerPheLys 272
Db 1427 AAAAGATAGCTAGAGATTAACCGCTTCCTTGCCCAAGAGTCTGAAGATTTCTTATAAA 1486
Qy 273 ArgLysArgPheLeuIleLysLeuArgProAspAlaAsnSerAlaTyGlnAspThrLeu 292

```
Db 1487 CGTAGTACGCTTTTCATCATCAATTCGGCTCGAGAGCAAGCAGCTATGAAGACCATC 1546
Qy 293 GlnPheLeuNeuEclAserArgAserPheCysValSerPheTyrValIleCysValGluHis 312
Db 1547 GGATTCGAACTCCCGATTCCGAGCAGCTGAAGAAATTAAGGAAAGCTGTGTGAACAT 1606
Qy 313 HisAlaPhePheArgLeuPheGluGluPheProPheProPheProValLeuPheSer 332
Db 1607 CACACGTTTTCACATTTG--ACATCTACAGACACCATTCACAAAGCAAAATTTCTTGCG 1663
Qy 333 ArgGlySerSerPheArgPheSerGlyArgThrGlnValLeuAspTyrValIys 352
Db 1664 CTAGGATCCAAATTCGATACAGTGGCCGACCTCA-- 1699
Qy 353 GluGlyGlnHisValValGlnPheGluArgHisSerLysIleHisSerIleArg 372
Db 1699 ----- 1699
Qy 373 SerLeuAlaSerGlnProThrGluLeuAsnSerGluValLeuGlnSerGlnGlnSer 392
Db 1700 -----GCTCAGACCGAGGAGCT 1717
Qy 393 ThrSerLeuThrPheGlyGluGlyAlaGluSerProGlyGlnSerCysArgArgGly 412
Db 1718 AGTGTCTTAATTT----- 1729
Qy 413 LysGluProLysValSerAlaGlyGluProGlySerHisProSerProAlaProArgArg 432
Db 1730 -----GACAGCGCTGCCCCACACTTC 1750
Qy 433 SerProAlaGlnValLeuGlnAla-----AspGlyAlaAlaSerAlaProThr 448
Db 1751 GAGCGTACAGCAAGTAAACGGCGCTCCGAGCCTCGATGAGACAGAGCTGTC----- 1804
Qy 449 GluGluGluGluGluValValAspArgThrGlnGlnSerLysProGlnProGln 468
Db 1805 -----GATTCGGCAGACCGAAGCTCCG 1828
Qy 469 ProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAsnSerGln 488
Db 1829 CCC-----ACTTCGACACCTGCCCATTTACTCTCAGGTCAGGTT--GCAGAA 1870
Qy 489 GlyGlyVal-----AlaProAlaAsnValThrLeuSerProAlaLeuSerProAspThr 506
Db 1871 GGTGGCGCTCTAGATGCTCTGCTAAACAAACAGTGTCTCTAAAGCAGACAGAGAAACA 1930
Qy 507 LysGlnAlaSerProLeuIleSerProLeuLeuAsnArgGlnAlaCysProArgThrAsp 526
Db 1931 GTGAAGCTGAAGTGAAGAAAGAAAGAGACGACCTGAGTGAAGCTGAGCCAGACCA 1990
Qy 527 AspGluAspGluGlyArgGlyArgPheProThrAspLysValTyrPheIleAlaLys 546
Db 1991 GAAGCATGAGAAAGAAAGAGAAAGACTAGATGAGTGAAGCAATTTATATC----- 2041
Qy 547 GluValSerThrGlnArgGlnThrTyrLeuLysAsnArgGlnValIleThrSerTyrPhe 566
Db 2042 -----AGACATACGCAATTTAATGTTGAGGATTTAGCAAG-----AGT 2080
Qy 567 GlnSerThrValSerLysGluAspAlaSerProGluAlaLeuLysSerLeuIlePhePro 586
Db 2081 CAAAGAGAGATCAAAAACATCATGCGACGATCACTAGTACGTAAGAAAG----- 2128
Qy 587 AsnPheGluProLeuHisValPheHisThrAsnPheLeuLysGluIle--GluGlnArg 605
Db 2129 -----AATTTCATGAGTCTGTAACAGAACCAAG 2158
Qy 606 LeuAlaLeuTyrGluGlyArg 612
Db 2159 CTTAGTGAATGGATAAACGC 2179
RESULT 4
US-09-664-958-5
```

```
/ Sequence 5, Application US/09664958
/ Patent No. 6916912
/ GENERAL INFORMATION:
/ APPLICANT: Trakhe, Ilya
/ APPLICANT: Canfield, Robert
/ APPLICANT: Kalantarov, Gary
/ APPLICANT: Rudchenko, Sergei
/ TITLE OF INVENTION: No. 6916912el Tumor-Associated Marker
/ FILE REFERENCE: 0575/60240
/ CURRENT APPLICATION NUMBER: US/09/664,958
/ NUMBER OF SEQ ID NOS: 28
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 5
/ LENGTH: 6263
/ TYPE: DNA
/ ORGANISM: Human
/ US-09-664-958-5

Alignment Scores:
Pred. No.: 1,12e-56 Length: 6263
Score: 712.50 Matches: 239
Percent Similarity: 41.5% Conservative: 99
Best Local Similarity: 29.3% Mismatches: 281
Query Match: 13.0% Indels: 198
DB: 3 Gaps: 31

US-09-555-342b-2 (1-1045) x US-09-664-958-5 (1-6263)
Qy 18 GluAsnSerGlyIleSer-----ThrLeuGluArgGlyGlnLysProProProThr 34
Db 369 GAGGCGAGTGGCTTTTCGAGAGAGACCGCCAGCAAGCCGAAATGCGCCAGAG 428
Qy 35 ProSerGlyLysLeuValSer-----IleLysIleGlnMetLeuAspThrGlnGlu 52
Db 429 ATTGCCAAGAAATACAAAGACTGTCATCTGCGGCTCACTGTGTTGATGCTGGAGTAT 488
Qy 53 AlaPheGluValProGlnAlaArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHis 72
Db 489 GAGTGTAGGTGAGAGAAACATGCGCGGAGCCAGTGTCTTTTAACTCGTCTGTGAAC 548
Qy 73 LeuAsnLeuValGluLysAspTyrPheGlyLeuGluPheProAspHisValIleThr 92
Db 549 CTGACCTCTTAGAAGAAAGACTACTTGGCTGACTCTTGATGCTGACAGCCAGAG 608
Qy 93 ValTyrLeuAspLeuLeuLysProIleValLysGlnIleArg--ArgProLysHisVal 111
Db 609 AACTGGCTGAGACCCCTCCAAAGAGATCAAGAACAGATCCGAGTAGACCCCTGAAATTTT 668
Qy 112 ValValLysPheValValLysPhePheProProAspHisThrGlnLeuGlnGluLeu 131
Db 669 GCC-----TTCACAGTCAAGTTTCTACCCGCTGATCTGCGCCAGCTGACAGAAACATC 722
Qy 132 ThrArgTyrLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCys 151
Db 723 ACAAGATACACTGCTGCGCTGAGCTGCGGAGCAGATATCAACGCGCGCGCTGATGC 782
Qy 152 AsnAspThrSerAlaLeuLeuLysIleSerHisIleValGlnSerGluIleGlyAspPhe 171
Db 783 TCTTTGTCAACGATGCCCTACTGAGGCTCTGACGCTGTGAGGCTGAGCTGAGTGTAT 842
Qy 172 AspGluAlaLeuAspArgHis-----LeuAlaLysAsnLysTyrIlePro 187
Db 843 GATGCT-----GAGAGCATGAGGCGCAATGATGTCAGCAAGCTCCGCTGCGCCCT 893
Qy 188 GlnGlnAsp--AlaLeuGluAspLysIleValGluPheHisValAsnHisIleGlyGln 206
Db 894 AACCAAGCCCGGAGAGCTGAGAGAGAGATCATGAGTGAATGAACATAGAGGAGATG 953
Qy 207 ThrProAlaGluSerAspPheGlnLeuGlnIleAlaArgArgGluLeuMetTyrGly 226
Db 954 ACCCGGAGAGACAGAAATCACTTCTTAGAAGATCCAAAGAAAGCTTTCATGATCGGA 1013
```

```

Qy 227 lIeaRgluEuhIaPfoAlaLyAaPArGluGlyThrLyIleuValaValaAaPn 246
Db 1014 GTAGACTGTCAGCATGCAAGAGCATGTCAGGCGATGACATCATGTTAGCGCTTTGGCC 1073
Qy 247 ThrGlyIleuValPheGlnGlyPheThrLyIleuAlaPheAenThrAlaVal 266
Db 1074 AATGGCTGCTGCTCATACCGGAGACGGGCTGAGATCAACGCTTTGCTGGCCCAAGATC 1133
Qy 267 ArgLyLeuSerPheLyAaGlyAaGlyPheLeuIleuLyLeuAlaPheAaSer 286
Db 1134 CTCAGAGTCTCTCTCAAGAGAGTAACTTTATATCAAGATCGAGCTGGGAGATAG 1193
Qy 287 AlaTyrAlaAaPThrLeuGluPheLeuMetAlaSerAaPheCysLySerPheTyr 306
Db 1194 CAATTGAGAGCACAATGGCTTTAAGCTCCAAACACCGGTCAGCAAGACATGTGG 1253
Qy 307 LyuIleCysValaGluIleValaAlaPhePheAaGlyLeuPheGluPhePro 326
Db 1254 AAGGTGCAATCAGACATCATATCTTCGGCTGTG---TCCCTGAGCCCAACC 1310
Qy 327 LyuPheValLeuPheSerAaGlySerSerPheAaPheSerGlyAaThrGlnLy 346
Db 1311 AAGGGCTCTGTGTG---ATGGCTCCAGATTCGGGTACAGTGGAGACCCAGCAAG 1367
Qy 347 ValLeuAaPThrValLyGluGlyGlyIleuValaGlnPheGluAaGlyAaSer 366
Db 1368 ACTCGCAGGCGCAGCGGCTCATGACCGGCTGACCCCTTTGAGCTTGTCCAGC 1427
Qy 367 LyuIleHisSerIle---ArgSerLeuAlaSerGlnProThrGluLeuAaSerGlyVal 385
Db 1428 AAGCGACACACATGTCCTCCAGCTTGATGAGACAGATTCCTCCGCCCACTGGTGC 1487
Qy 386 LeuGluGln-SerGlnGlnSerThrSerLeuThrPheGlyGlyValaGluSerPro 405
Db 1488 AGCGAGACCATATGATGAGGGCTGACGGTGAACAGCGGATAGATGGCGAGTCTGG 1547
Qy 405 YGlyGlnSerCysArg----- 410
Db 1548 GGGCAACGGTCAGAGGCTGAGAGAGAGTCAAGCTCAACCAAGATCAAGAGACTA 1607
Qy 411 ---ArgGlyLyGluPro-----LysVa 417
Db 1608 AAGCCGAGAGCAAAACACGCGGAGACACAGACAGAGTCTTAAGACAGACAGAGAT 1667
Qy 417 lSer-----AlaGlyLupProGlySerHisPProSerProAlaProAaGlySerProAl 435
Db 1668 GTCTTGCTGACAGCACAGGCCAGCATATGAGCTCAAAAGACCTTGAGAGGCCAAC 1727
Qy 435 agLyAaPn-----LysGlnAlaAaPglyAlaAlaSerAlaProThrGlu 450
Db 1728 AGCAAACTCATCAACGGGATCGAGACTGGGAACGGGAGCGAGGCTGCC----- 1779
Qy 450 uGluGluGluValValyAaPArGThrGlnGlnSerLyPheGlnProGlnProSe 470
Db 1780 -----CTTCCCGCGCTCC-CCCTC 1798
Qy 470 rThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAaSerGlnLy 490
Db 1799 CCCC-----AAGGGCACCCCTGAG-----AAAGCCATAGAGAGCAGG 1837
Qy 490 yValAlaProAlaAaValThrLeuSerProAaLeuSerProAaPThrLySerGlnAlaSe 510
Db 1838 GCTGAGGAGAGGCTCC-----GAGGAGAAAGTCAAAACACACAGTCCCGGGCC-- 1886
Qy 510 rProLeuIleSerProLeuLeuAaPheGlnAlaCysPProAaGThrAaP----- 526
Db 1887 -----CCAGAGAGTGAACACAGCGCATGA 1909
Qy 527 -AaPglyAaPglyGlyAaGlyAaGlyAaPheProThrAaPlyAlaAlaTyrPheIleAlaLy 546
Db 1910 GGAACCGAGACAGAGAGGAGACAGCGGTCTCTGAGAGACAAACACTGGCCATTGAGCG 1969
Qy 546 aGluValaSerThrThrGluAaGThrTyLeuLyAaPLeuGlu----- 560

```

```

Db 1970 CAAGTCTCCAGACATCAACGGTCAGCTCTAGCTGAGGCTGAGGCTGAGGACTTCAAC 2029
Qy 561 -ValIleThrSerTrpPheGlnSerThrValSerLyGluAaP-----AlaMetPr 577
Db 2030 GGTCAATGGTGAATCAATGAGCAGCGCTTC-----GAAAGCTTCTCCCGAGCTGAC 2083
Qy 577 oGluAla-----LeuLySerLeuIlePhe----- 585
Db 2084 TGAAGTCGACCGGAGCAAAAGGAGCTCGGACACTGAGGCGCTGCTGTCTCCGGGATCT 2143
Qy 585 ----- 585
Db 2144 CAACAAGGGGGCCCCAGCCAGATGATGATGCTGGGGCATTTGAGAGACGCGGATG 2203
Qy 586 -----ProAaPheGluPheLeuIleValPheIleThrAa 597
Db 2204 AAGGGCTGCTCCACCCGGATATGCGCCAGTTTGAAGCCCGTG---AAACAGAAACAT 2260
Qy 597 nPheLeu-----LyGluIleGluGlnAaGlyLeuAlaLeuTyr----- 609
Db 2261 GACTGTCAGAGTCTGGCCATTAGAAAGATGAGACCGGAGCGGCTACTGACAGCAG 2320
Qy 610 ---GluGlyAaGSerAaPAlaGlnIleAaGArTyGlnAaGlyIleGlyAaPValMetIle 628
Db 2321 AGTCTCCGATGATGATTAACACACAGAGGTGATGG- GAGTCCCTCAGTGGGAGGAGT 2379
Qy 628 uLyAaAenIleGlnGlyMetLyHisLeuAlaAlaHisLeuTrrLyHisSerGlyAlaLe 648
Db 2380 TCATAGCAACCACTCCCTCATTCACCGAGAGACATATGACACACATGAGAAACATTC 2439
Qy 648 uGluAlaLeuGluAaGly----- 654
Db 2440 TCAGTCCGGGAAGGGGGAGCTGCCATGATCCAGGCCCAAGACGTCGCGCAGGAAA 2499
Qy 655 ---IleLySerSerAaGlyAaGlyAaPheCysAaGArPheGlyLeuGlnLyVa 673
Db 2500 TCCGTTCTCTTCTCCGATCATCGGGAAGA---TGTCTCCACAGACACTACGGCGCA 2556
Qy 673 lCysTyLeuProLeuAaPThrPheLeuLeuAaPProLeuHisValGlyMetHisTyLy 693
Db 2557 CTGGCGAAACCTCTCAACTC----- 2578
Qy 693 aGlnValLeuGluAaGlyLeuCysLyHisAaPProSerHis 707
Db 2579 -----CACCAACACCATGTCAAC 2596

RESULT 5
US-09-566-921-57
/ Sequence 57, Application US/09566921
/ Patent No. 6682888
GENERAL INFORMATION:
/ APPLICANT: Loring, Jeanne F.
/ APPLICANT: Tingley, Debora W.
/ APPLICANT: Edwards, Carla M.
/ TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
/ FILE REFERENCE: PA-0024 US
/ CURRENT APPLICATION NUMBER: US/09/566,921
/ NUMBER OF SEQ ID NOS: 138
/ SOFTWARE: PERL Program
/ SEQ ID NO 57
/ LENGTH: 6268
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc. feature
/ OTHER INFORMATION: Incyte ID No. 6682888 453592.3
/ NAME/KEY: unsure
/ LOCATION: 2956, 4230-4252, 4825-4887
/ OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-57

```

Alignment Scores:

Pred. No.: 1,12e-56 Length: 6268
 Score: 712.50 Matches: 239
 Percent Similarity: 41.54 Conserved: 59
 Best Local Similarity: 29.34 Mismatches: 281
 Query Match: 13.04 Indels: 198
 DB: 3 Gaps: 31

US-09-555-342B-2 (1-1045) x US-09-566-921-57 (1-6268)

QY 18 GluanserGlyIleSer-----ThrLeuGluArgGlyGlnLysProProThr 34
 Db GAGGCCGATGGCTTTGCGAGAGGACACGCCAGACAGCCAGAAATCCCGAGAG 428
 QY 35 ProserGlyLysLeuValSer-----IleYsIleGlnMetLeuAspThrGlnGlu 52
 Db ATTGCCAAGAAATCAAGAGTGCATCTGCCGGCTCATCTGTGAATGCCGAGAT 488
 QY 53 AlaPheGluValProGlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHis 72
 Db GAGGTGAGGTGGAGAAACATGCCGGGGCCAGGTGCTGTTGACCTGTGTGAACAC 548
 QY 73 LeuAsnLeuValGlnGlyAspTyrPheGlyLeuGluPheProAspHisLysLysIleThr 92
 Db CTCAACCTCTTGAAGAAGGACTACTTGGCTGACCTTGTGATGTGACAGCCAGAG 608
 QY 93 ValTTrpLeuAspLeuLysProIleValLysGlnIleArg--ArgProLysHisVal 111
 Db 609 AACGTGGCTGGACCCCTCCAGAGAGATCAAGAGATCCGAGATAGCCCTGGAATTT 668
 QY 112 ValValLysPheValValLysPhePheProAspHisThrGlnLeuGlnGluLeu 131
 Db 669 GCC-----TTCAAGTCAAGTTTACCCGCTGATCTCGCCAGCTGACAGAAAGATC 722
 QY 132 ThrArgTyrLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCys 151
 Db ACAAGATATCTACTGTGCTGCTGACGTGGGAGACATCATCAGGGCCGCTGCATGC 782
 QY 152 AsnAspHisSerAlaAlaLeuLeuIleSerHisIleValGlnSerGlnIleGlyAspPhe 171
 Db 783 TCTTTGTACGATGCGCTTACTGGCTCTTACGCTGTGAGGCTGAGCTGGAGCAT 842
 QY 172 AspGluAlaLeuAspArgGluHis-----LeuAlaLysAsnLysIlePro 187
 Db 843 GATGCT-----GAGAGCATGTGGGCAATATGTACAGCAGCTCCGCTGCCCT 893
 QY 188 GlnGlnAsp--AlaLeuGluAspLysIleValGlnPheHisIleAsnHisIleGlyGln 206
 Db 894 AACCAAGCCGGAGCTGGAGAGAGATCATGAGGTGCATAGACATATAGGGGGATG 953
 QY 207 ThrProAlaGluSerAspPheGlnLeuLeuGlnIleAlaArgArgLeuGluMetTyrGly 226
 Db 954 ACCCGGAGAGAGCAAAATCCACTTCTTACAGATGCAAGAAAGCTTTCATGTACGGA 1013
 QY 227 IleArgLeuHisProAlaLysAspArgGlnGlyThrLysIleAsnLysAlaValAlaAsn 246
 Db 1014 GTAACTGTGACCACTGCCAGAGACTGTAGGGCATCCAGATCATGTTAGGGCTTTGGCC 1073
 QY 247 ThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAlaPheAsnTrrAlaLysVal 266
 Db 1074 AATGGCTGCTCATCTACCCGGGACCGGCTGAGATCAACGCTTTGCTGGCCCAAGATC 1133
 QY 267 ArgLysLeuSerPheLysValGlyLysArgPheLeuIleLysLeuArgProAspAlaAsnSer 286
 Db 1134 CTCAGATCTCTCTCAAGAGAGATTAATCTTATATCAAGATCCGCGCTGGGAGATATGAG 1193
 QY 287 AlaTyrGlnAspTrrLeuGluPheLeuMetAlaSerArgAspPheCysLysSerPheTrr 306
 Db 1194 CAATTGTAGAGCACAATTTGGCTTTAAGCTCCCAAAACACCGGTGACGCCAAGAGCTGTGG 1253
 QY 307 LysIleCysValGlnHisIleAlaPhePheArgLeuPheGlnGluProLysProLysPro 326
 Db 1254 AAGGTCTGATCGAGCATCATATCTTCCGGCTGGTG--TCCCTTGAGCCCCCAGCCC 1310

QY 327 LysProValLeuPheSerArgGlySerSerPheArgPheSerGlyArgThrGlnLysGln 346
 Db 1311 AAGGGCTCTTCGTGTG--ATGGGCTCCAAAGTTCCGGTACGTGGAGAGACCCAGGCAAG 1367
 QY 347 ValLeuAspTyrValLysGlnGlyGlnHisLysLysValGlnPheGluArgLysIleSer 366
 Db 1368 ACTGCCAGGCGCACGGCCCTCATTTGACCGGCTTCACCCCTTCTTGAAGCGTTTCCAGC 1427
 QY 367 LysIleHisSerIle--ArgSerLeuAlaSerGlnProThrGluLeuAsnSerGluVal 385
 Db 1428 AAGCGTACACCATGTCTCCGAGCCTTGATGACAGAGTTCTCCGCCAGCTCCGCTC 1487
 QY 386 LeuGlnGln--SerGlnLysSerThrSerLeuThrPheGlyGlnGlyAlaGluSerProGly 405
 Db 1488 AGCGAAGAACATGATGTGACGGGCTGACGGTGAACAGCGGATGAGATGGCGAGTCTGGG 1547
 QY 405 YGlyGlnSerCysArg----- 410
 Db 1548 GGGCAAGGTCAAGAGCTGAGAGGAGAGAGTCAAGATCCAAAGATCAAGAGACTA 1607
 QY 411 --ArgGlyLysGluPro-----LysVal 417
 Db 1608 AAGCCGAGAGAGAAACACAGCCGAGACACAGAGAGTTCTTGAACAAAGCAGAGAT 1667
 QY 417 Lser-----AlaGlyLysProGlySerHisProSerProAlaProArgHisSerProAl 435
 Db 1668 GTCTTGCTGAAGCACCGAGCCAGCATATGAGCTCAAAAGACCTTGAAAGAGCCCAAC 1727
 QY 435 ArgLysAsn-----LysGlnAlaAspGlyAlaAlaSerAlaProThrGlnGly 450
 Db 1728 AGCAAACTCATCCACCGGAGATGACAGTGGGAACGGAGGCGAGCGCT----- 1779
 QY 450 uGlnGlnGluValValLysAspArgThrGlnGlnSerLysProGlnProGlnProse 470
 Db 1780 -----CTTCCCGGCTCC-CCCTC 1798
 QY 470 rThnGlySerLeuThrGlySerProHisLeuSerGlnLeuSerValAsnSerGlnGly 490
 Db 1799 CCCC-----AAGGCAACCCCTGAG-----AAAGCATGAGAGAGCAGG 1837
 QY 490 YValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAspThrLysGlnAlaSer 510
 Db 1838 GCTGAGGAGAGGCTCC-----GAGAGAAAGTCAAAACACACAGCTCCCGGGCC-- 1886
 QY 510 rProLeuIleSerProLeuLeuAsnArgGlnAlaCysProArgThrAsp----- 526
 Db 1887 -----CCAGAGGTGACACAGCGGATGA 1909
 QY 527 -AspGluAspGlnGlyArgArgLysArgPheProThrAspLysAlaLysPheIleAlaLys 546
 Db 1910 GGAACAGAGACAGAGAGGAGACAGCGGTCTTGAAGAGAACACCATCGGCATTTGAGCG 1969
 QY 546 ArgLysAspThrThrGlnLysArgThrTyrLysLysAspLeuGlu----- 560
 Db 1970 CAAGTGTCCAGCATTCAGGTCTAGCTTACGTCTGAGGCTGAGGTGAGTTCAC 2029
 QY 561 -ValIleThrSerThrPheGlnSerThrValSerLysGluAsp-----AlaMetPyr 577
 Db 2030 GGTCAATTGTGATACATCATGAGCGGCTTC-----GAAGACTTCTCCCGAGCGCTGCC 2083
 QY 577 OGluAla-----LeuLysSerLeuIlePhe----- 585
 Db 2084 TGAAGCTGACCGGAGCAAAAGCACTTCGACACTGAGGGCTGTCTGTTCTCCGGGATCT 2143
 QY 585 ----- 585
 Db 2144 CAACAAGGGGGCCCCAGCAGATGATGAGTGTGGGGGATTTGAGAGACCCCGGATCG 2203
 QY 586 -----ProAsnPheGluProLeuHisLysPheHisThrAs 597
 Db 2204 AAGGGCTGTGTCACACCCCGGATGATGCCCGAGTTTGAGCCCGTG--AAAACAGAAACCAT 2260

Qy 597 npheluu-----lysgluilegluigluargleuialeuTrp----- 609
Db 2261 GACTGTGACAGTCTGGCCATAGTAAAGAAAGATTGAGCCGAGCCGCTACTGACAGACCAG 2320
Qy 610 ----GluGlyArgSerAsnAlaGlnIleArgAspTyrGlnArgIleGlyAspValMetCLe 628
Db 2321 AGTCTCGCATGATGATACACCCAGCGATTGATGG--GAGTGGCTTCAGTGGGAGGAGT 2379
Qy 628 uLyAsnIleGlnGlyMetCLeHisLeuAlaIleHisLeuTrpLysHisSerGluAlaLe 648
Db 2380 TCATAGACACCATCTCCATCATCACGAGACCATATGACACCATGAAACAGTCTC 2439
Qy 648 uGluAlaLeuGluAsnGly----- 654
Db 2440 TCAAGTCCGGAGAGGGGCGAGCTGCATATCCAGGCCACAGACGCTGCCACGGA 2499
Qy 655 ----IleLysSerSerArgArgLeuGluAsnRheCysArgAspRheGluLeuGlnLysVa 673
Db 2500 TCCGTTCTCTTTCTCCATCATCGGGAAGA---TGTCTCACACGACCATACGCGCCA 2556
Qy 673 lCysTyrLeuProLeuAsnThrRheLeuLeuArgProLeuHisArgLeuMetHisTyrLys 693
Db 2557 CTGGGAAACCTCTTCACTC----- 2578
Qy 693 vGlnValLeuGluArgLeuCysLysHisHisProProSerHis 707
Db 2579 -----CACCAACAACCATGTCTAC 2596

RESULT 6
US-09-799-451-201Sequence 201, Application US/09799451
Patent No. 6783969

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aiding J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yunding
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6783969el Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799, 451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: PL_FL_genes Version 2.0
SEQ ID NO 201
LENGTH: 3398
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (174) .. (2234)
US-09-799-451-201

Alignment Scores:

Pred. No.: 7.2e-49
Score: 626.00
Percent Similarity: 44.6%
Best Local Similarity: 30.1%
Query Match: 11.5%

Length: 3398
Matches: 191
Conservative: 92
Mismatch: 221
Indels: 130

DB: 3 Gaps: 22
US-09-555-342B-2 (1-1045) x US-09-799-451-201 (1-3398)
Qy 31 ProProProThrProSerGlyLysLeuValSerIleLysIleGlnMetLeuAspRhr 50
Db 273 CCGCAGCTGGAGATTCTAAGTTCATACATCAGCTGCGGCTGCCCTTCTGGATGGTACT 332
Qy 51 GlnGluAlaPheGluValAlaProGlnArgAlaProGlyLysValLeuAspAlaValCys 70
Db 333 GATGTTAGTGTGACTCTGCCAAAAAGCCAAAGCCAAAGAGTTGTTGATTCAGATTAG 392
Qy 71 AsnHisLeuAsnLeuValGluLysArgTyrPheGlyLeuGlnPheProAspHisLysLys 90
Db 393 TACCACTGGACCTGATTGAAAGCGACTATTTGGCTGAGATTATGATTGACAGACA 452
Qy 91 lIethValITrPLeuAspLeuLysProIleValLysGlnIleArgArgProLysHis 110
Db 453 GTAGCAATTTGGTGTGATGTACAAAACATCAAAAACAAAGTAAAAATTGGTTCACCC 512
Qy 111 ValValValLysRheValValLysRheRheProProAspHisIethGlnLeuGlnGlu 130
Db 513 TATTTCTGCATCTTGCAGTTAAGTTTATTTCTCTCAAAACCAATACCTTCTGTAGAG 572
Qy 131 LeuThrArgTyrLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnIlyArgLeuThr 150
Db 573 CTAAACCGGATTTATTTGTTCTTCAAGTAAACAAAGATATTTCTCAGTGAATAATGAC 632
Qy 151 CysAsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAsp 170
Db 633 TGTCCCTTGATACAGCAGTGCATGCAATTGGCAGCTTATATCTCAAGCTCAACTGGTGC 692
Qy 171 PheAspGluAlaLeuAspArgLysHis-----LeuAlaLysAsnLysTyrIle 186
Db 693 TATGATCTTGGCT-----GAGCATAGTCTCTGAACCTGTCTCAGAGTTCAGATTCTG 743
Qy 187 Pro---GlnGlnAspAlaLeuGluAspLysIleValGluPheHisIleAsnHisIleGly 205
Db 744 CTTATTCAGACTGAGATGAGATGAGATGCGCTATTTTGAAGAAATGGAAGAAATGACAGGT 803
Qy 206 GlnThrProAlaGlnSerAspRheGlnLeuGluIleAlaArgArgLeuGlnMetTyr 225
Db 804 CAACACAGCAGCAGGCTGAAACCAATTATCTGAATTAAGCCAAAGCGCTAGCAAAAGTAT 863
Qy 226 GlyIleArgLeuHisProAlaLysAspArgGlnGlyThrLysIleAsnAlaValAla 245
Db 864 GGGGTGATATGATGATGTGTCAAGGCTAGAGATGGGAATGACTATATGTTGGGACTAACA 923
Qy 246 AsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAlaPheAsnITrAlaLys 265
Db 924 CCAACAGAGATCCTTGTGTTTGAAGAGATACCAAAATGGCTTATTTTGGCCGGAAG 983
Qy 266 ValArgLysLeuSerPheLysArgLysArgPhe---LeuIleLysLeuArgProAla 284
Db 984 ATAAACAGATGGATTTTGAAGAAATTAATTAACCTGTGTGTGAAGATGTGAT 1043
Qy 285 AsnSerAlaTyrGlnAsp---ThrLeuGlnPheLeuMetAlaSerArgAspRheCysLys 303
Db 1044 CAGGCAAAAGAACAGAAACATCATTTGCTTTAAGCTGATCATCAAAAGCATCAAA 1103
Qy 304 SerPheTrpLysIleCysValGlnHisValaPhePheArgLeuPheGlnGluPro--- 322
Db 1104 CATTTATGAAATATGCTGTGAGCATCATGCTTCTTCCGCTT---CGAGGCCCGGTC 1160
Qy 323 ---LysProLysProLysProValLeuPheSerAlaGlySerSerPheArgPheSerGly 341
Db 1161 CAAGAAGTTCTCATCAGATCAGAGATTTATTCAGATGATCAAGATTAATGATATGAG 1220
Qy 342 ArgThrGlnLysGlnValLeuAspTyrValLysGlnGlyLysValGlnPhe 361
Db 1221 AAAACAGATATCAGACCAAAACCAATTA---GCAAGAAGATCAACATCTTT 1274
Qy 362 GluArgLysHisSerLysIleHisSerIleArgSerLeu-----AlaSer 376


```

Db      1275 GAAAGAGGCCAGAGAAAGCATATCTAGACAGACTCTACAAATGAAGCATGCTACAA 1334
Qy      377 GlnProthrGluLeuAsnSerGluValLeuGlnInserGlnInserThrSerLeuThr 396
Db      1335 AAACCTGAGAACTT-----AGTGTTCACAAATATGTTTCGACC 1373
Qy      397 PheGlyGluGlyAlaGluSerProGlyGlyInserCysArgArgGlyLysGluProLys 416
Db      1374 CAAGTAATGGCTCCCAACAGAGCTTGGGGAGTAGATGTCTGCTGAGTCTTCC 1433
Qy      417 ValSerAlaGlyGluProGlySerHisProSerProAlaProArgArgSerProAlaGly 436
Db      1434 ATTTC----- 1439
Qy      437 AsnLysGlnAlaAspGlyAlaLysAlaSerAlaProThrGlnGluGlnGluValLys 456
Db      1440 -----TCTGCTCTGTGCGCATGAGATAGAG----- 1466
Qy      457 AspArgThrGlnInserLysProGlnProGln-----SerProLysSerGluLeuSerValAsnSer 487
Db      1467 -----AATCTTCACAGAGCTCTGGAACAGACAGCATGACAGAAATGC 1511
Qy      469 ---ProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAsnSer 487
Db      1512 ATTCTCTGATATGATGATTGCTGATAGCCACAGACTTATGGAACAGACTT----- 1565
Qy      488 GlnGlyGlyValAlaProAlaAsnValThrLeuSer----- 499
Db      1566 ---GGTGAATGTAATGGGGCATCTGACACTATGAAACATCCAAAGCAGTGAATGACGTT 1622
Qy      500 -----ProAsnLeuSer-----Pro 504
Db      1623 AATGTAGCCACAGAGCTTCCGGATAGGGGAAGCTGAAGTTGAATGACAGACTTAAAA 1682
Qy      505 AspThrLysGlnAla-----SerProLeuLysSerProLeu 516
Db      1683 GAGACCTCAGAGAGAGCTCAAAACAGCTTGAGATGAGAGAACACTCTCTTCTCTCCCTCGA 1742
Qy      517 LeuAsn-----AspGlnAlaCys 522
Db      1743 TCCAACATCGATGTTAATCAATAAACAGCAGAGAGAGATGTGAAGTTGACAGAAATATGC 1802
Qy      523 ProArgThrAspArgGluAspGluGlyArgArg---LysArgPheProThrAsp----- 539
Db      1803 CTTAATTAATGCTAGTAGAGCCAGAGATGTAATGATCATGAGATTCCTCTGACTTCAG 1862
Qy      540 -----LysAlaLysPheIleAlaLysGluValSerThrThrGluArgThrTyr 555
Db      1863 AGTAACATTTGAAAGGCTCAAGTAGAACAGTGCATAGGTTACAAAGAAAGATAGCTTA 1922
Qy      556 Leu-----LysAspLeuGluValIleThrSerTrpPheGlnSerThrValSerLysGlu 573
Db      1923 TTAAGTCATAAAATGCGCAATGTTCCAGGATGCTGCCAACAACAGTGCCTGTAAATGAG 1982
Qy      574 AspAlaMetPro-----GluAlaLeuLysSerLeuIle 584
Db      1983 AATAATGTGCCCTCCCAAGAGTCTCTGAGACTGTAG 2024

RESULT 7
US-09-906-779-3
; Sequence 3, Application US/09906779
; Patent No. 6770466
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: Human Protein Tyrosine Phosphatase Polynucleotides, Polypeptides,
; FILE REFERENCE: PT040P1
; CURRENT APPLICATION NUMBER: US/09/906,779
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US01/01563
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/176,306

```

```

; PRIOR FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2872
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-906-779-3

Alignment Scores:
Pred. No.: 5,7e-43 Length: 2872
Score: 563.00 Matches: 149
Percent Similarity: 50.8% Conservative: 82
Best Local Similarity: 32.7% Mismatches: 174
Query Match: 10.3% Indels: 51
Db: 3 Gaps: 12

US-09-555-342B-2 (1-1045) x US-09-906-779-3 (1-2872)
Qy      30 LysProProProThrProSerGlyLysLeuValSerIleLysIleGlnMetLeuAspAsp 49
Db      6 GAGCGCGCGCGCGCGCGCCAG-GCCACCTCTACTGCGCGCTTCTGCTGAGCGG 64
Qy      50 ThrGlnGluAlaPheGluValProGlnArgAlaProGlyLysValLeuLeuAspAlaVal 69
Db      65 ACCGAAATGAGCGCGTGAACCTGCCGAAACATGCCAAAGCCAGATTTGTTGATCAAGATT 124
Qy      70 CysAsnHisLeuAsnLeuValGlyLysPheGlyLeuGlnPheProAspHisLys 89
Db      125 GTGTACACATGGACCTTGTGGAAACAGATTACTTGGCTCCAGTTCTCGACTCTGCC 184
Qy      90 LysIleThrValTrpLeuAspLeuLeuLeuProIleValLysGlnIleArgArgProLys 109
Db      185 CAGGTTCCGACCTGGCTGGATCATGCGCAACCCATAAAAGAGATGAATTTGAACT 244
Qy      110 HisValValValLysPheValValLysPhePheProAspHisThrGlnLeuGlnGlu 129
Db      245 GCTTATGCTTACACTTTCAGATTAATACATTTCTTCAGAACCAACAACTTCGTGAG 304
Qy      130 GluLeuThrArgTyrLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeu 149
Db      305 GAGTTTCAAGAGTACCTGTTGTTTCAACTCAGCAGATCAATCTTCTGGAATAATGG 364
Qy      150 ThrCysAsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGlnInserGluIleGly 169
Db      365 AAATGCCCTTATGAACAAGCTGTGAATTAAGTCTCTGTCTACAAAGGAGGACTGGG 424
Qy      170 AspPheAspGluAlaLeuAspArgLysHis-----LeuAlaLysAsnLysTyr 185
Db      425 GAGTGCAGCTTCCA-----GAACACACACACAGCTGTGTGCTGAGTTCCGGTTTC 475
Qy      445 IleProGlnGln---AspAlaLeuGluAspLysIleValGluPheHisIle 204
Db      476 ATTCCAATGACACAGAAAGCAATGGAATTTGATTTCTTCACAGATGGAAGAGTGCAGG 535
Qy      205 GlyGlnThrProAlaGluSerAspPheGlnLeuLeuIleAlaArgGluLeuMet 224
Db      536 GGAAGAGCCCTGCGCAGGCGAACTCTCTATGTGAATTAAGGAAAGTGCCTGGAATG 595
Qy      225 TyrGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysIleAsnLeuAlaVal 244
Db      596 TATGGGCTAGACATGCACCTGTTCAGGGAGAGATGGCTGTGAATTTCTCTTGGACTG 655
Qy      245 AlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAlaPheAsnTrpAla 264
Db      656 ACCCGCAGACAGCATATTAATCTTGAAGAGCTACCAAAATAGGCTTATTTTGGCCT 715
Qy      265 LysValArgLysLeuSerPheLysArgLysArgPhe---LeuLysLeuArgProAsp 283
Db      716 AAAATTTCACAAATGGAATTTTAAAGCAAAATGACACTCGGTGCTGAGAGATAT 775
Qy      284 AlaAsnSerAlaTyrGlnAsp---ThrLeuGluPheLeuMetAlaSerArgAspPheCys 302

```

```
Db 776 GATCAGGAGCGTGAAGCAAGCACACCTTTGTGTTCCGTTTAAAGACAGTCCAGACCTGC 835
Qy 303 LysSerPheThrIleGlyValGluHisIleAlaPhePheArgLeuPheGluPro 322
Db 836 AAACACCTTTGGAAAGTGTGACGTTGACACACGACATTTCCGACTGCCGAGCCAGCA 895
Qy 323 LysProIlePheProIleProValLeuPheSerArg---GlySerSerPheArgPheSerGly 341
Db 896 AAGACGAAATCCATATAGATCCGACTTATCAGGCTGGCTCTGCTTCAGATTCAGTGGG 955
Qy 342 ArgThrGlnIleGlnValLeuAspTyrValLysGluGlyLys-----HisLysLysVal 359
Db 956 CGGACGAAATATCA-----GCTACACATGCTCCAGGTTTACGAAAGACCG 1003
Qy 360 GlnPheGluArgLysHisSerLysIleHisSerIleArgSer-----LeuAla 375
Db 1004 ACCTTTGAGAGAAAGCCTTATGAAAGCTTATCCATCCCGGAGACATTCACGTTCAAGCA 1063
Qy 376 SerGlnPro-----ThrGluLeuAsnSerGluValLeuGluGln 388
Db 1064 AGCAACCCAGTATAGACAGCCAGCTGCTCTTAAACAAATCCAGAAATCCATAATTAC 1123
Qy 389 SerGlnIleSerThrSerLeuThrPheGlyGluGlyAlaGluSerProGlyGlyGlnSer 408
Db 1124 CAGCTTCA----- 1132
Qy 409 CysArgArgGlyLysGluProLysValSerAlaGlyGluProGlySerHisPro---Ser 427
Db 1133 -----TATCATCTTAAATATCCATCCAGCCAGCCCGGTGGCATCTCTCACTCT 1180
Qy 428 ProAlaProArgArgSerProAlaGlyAsnLysGlnAlaAspGlyAlaAlaSerAlaPro 447
Db 1181 CCAATGTACAGCGCATCTTTCAGAGATGACAGATCGCATTTGAAGATCGGCGCAGTGA 1240
Qy 448 ThrGluGluGluGluValValLysAspArgThrGlnGlnSer 462
Db 1241 GATGACAGCCATTTGATTATGTCCAGACGACAGAAACGAAAGAAC 1285

RESULT 8
US-10-104-047-41
; Sequence 41, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 2493
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-41

Alignment Scores:
Pred. No.: 1,49e-42 Length: 2493
Score: 557.50 Matches: 200
Percent Similarity: 43.5% Conservative: 103
Best Local Similarity: 28.7% Mismatches: 248
Query Match: 10.2% Indels: 146
DB: 3 Gaps: 27

US-09-555-342B-2 (1-1045) x US-10-104-047-41 (1-2493)
Qy 8 ProThrProGly-----SerArgLeuGlnAlaProGlyLeuAsnSer 20
Db 163 CCTGGTCCAGAGGTCCGACGCGCGCGCTCGCTCCCGGCGGCGGAGAGATGCT 222
Qy 21 GlyIleSerThrLeuGluArgGlyGlnLysProProThrProSerGlyLysLeuVal 40
```

```
Db 223 GAG---CAGTTGATGAGCGGAGCAGCAGAGCTT-----GAGCGCGAGTA 267
Qy 41 -SerIleLysIleGlnMetLeuAspArgThrGlnGlnAlaPheGluValProGlnArgAl 60
Db 268 CAGCTGACCGGTGGCGCTCTGGAGCAGCAGATACACTGCACCATCCAGCAAGATGC 327
Qy 60 aProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGluLysArg 80
Db 328 CAAAGCCAGATCACTTTGACCTTTTGCCACCATCGAACCTTCTTGAGAAACATA 387
Qy 80 rPheGlyLeuGluPheProAspHisLysLysIleThrValTrrPLeuAspLeuLeuLysPr 100
Db 388 TTTGGTATCCGCTTGTAGACCCAGATAGACGCGCATGCTGGCTGAATTT----- 439
Qy 100 oIleValLysGlnIleArgArgProLysHisValValLysLysPheValValLysPhe 120
Db 440 -----ACAAAGTCTGTGTGAAGTTTA 462
Qy 120 eProProAspHisThrGlnLeuGlnGluGluLeuThrArgTyrLeuPheAlaLeuGlnVa 140
Db 463 TCTTGACAGACCTGCTGCTCTGAAAGAAATAACACAGTATTTAGTCTTCTGCAAT 522
Qy 140 LysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAlaAlaLeuLeuI 160
Db 523 CAAAGGATCTCTACATGCGCGACTCTCTGTAAACATGAGATGCTGCTTTTAC 582
Qy 160 eSerHisIleValGlnSerGluIleGlyAspPheArgLysAlaLeuAspArgLys---His 179
Db 583 AGCTTACATCTTCAAGCGGAATGGGATTAATGCTCAGAGAAACACCTGAAGGCTA 642
Qy 179 sLeuAlaLysAsnLysTyrIleProGln---GlnAspAlaLeuGluAspLysIleValG 198
Db 643 CAGCTCCAAAGTTCCAGTTTTCCTTAAACATTCAAGAAAGTGGAAAGAAATTCCTGA 702
Qy 198 uPheHisHisAsnHisIle---GlyGlnThrProAlaGluSerAspPheGlnLeuLeuG 217
Db 703 GATTCAACAGACGAACTGATGTGTCAACACACAGACATCATGAGCTGAATCTTAAAG 762
Qy 217 uIleAlaArgArgLeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGluG 237
Db 763 AAAAGCACACATTTGAAACATATGAGTGAATCTCCACCAATGTATAGAGAGTGTGACG 822
Qy 237 yThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLy 257
Db 823 AATGCTGATTTCTGGCTTCACTCCCTTTGGTTGTGTTCTTCAAGAAACAAAGAG 882
Qy 257 sIleAsnAlaPheAsnThrAlaLysValArgLysLeuSerPheLysValGlyArgPheLe 277
Db 883 GGTCCACTTCATTAATGAAATGAGGTGACCAAGCTGAATTTGAAGAAAGACTTTCTA 942
Qy 277 uIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAl 297
Db 943 TTTATACGTA-----AGTCAGAAAGAGAAAGAAATTAATTTCAATATTTTCTCC 996
Qy 297 aSerArgAspPheCysLysSerPheThrLysIleCysValGluHisIleAlaPhePheArg 317
Db 997 AACTCTGAAGGCTTGAAGCACCTCTGAAATGTGAATCGAAGAACCAAGCCTTCAACA 1056
Qy 317 GluPheGluGluProLysProLysArgProVal-----LeuPheSerArg 333
Db 1057 GCTG-----GAGAAGTCAGAGCCAGTCCGACAGTGTCCACAGCAATTTATTTCTTAA 1110
Qy 333 gGlySerSerPheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysG 353
Db 1111 AGGAGCCGCTTCCATACAGTGGCGGAGTTCCAAAGAAAGATCATAGAA----- 1159
Qy 353 uGlyGlyHisLysLysValGlnPheGluArgLysHisSerLysIleHis----- 369
Db 1160 -----TCAAGTCTAATATCAACGAGGACCGCAAGAAATACACAGACGAGGAT 1209
Qy 370 -----SerIleArgSerLeuAlaSerGlnProThrGluLeuAsnSerGluVa 385
```

Db 1210 GGTCCACCGCCGAGCTGCTCCTCCATTAACCATAGCCCAAG-----CTGAGCAGGCT 1265

Qy 385 lIeugluInsergInserThrSerleuthrPheglYgluYalagIuser----- 403

Db 1264 CCCCAGACCCGCAAGAGGCTTTCATCTCCATCATGGAAGGCTTAGAGTCTTACG 1323

Qy 404 -----Progl 405

Db 1324 GGACAGTCCCATTCACACAGTGCCTTCCATTCAGTGGGACACCTTCTGCTCA 1383

Qy 405 YglYglInserCybATgATgglYLygluIuproYvalSer-----AlaGl 420

Db 1384 CGTGAAAGCAGCCGCAAGATAGCAATGAGGAGTAgCTGTGATTCAGACAGAGCTTA 1443

Qy 420 YgluProglYserThIs-----ProSerProAlaProArgrArSer-----ProAl 435

Db 1444 CAGCCTTCAGACAGCGTGCCTGCCACCCCTGTGCTGAGCAGCCTCGAGCTGATTT 1503

Qy 435 AGLYaenLYseGlAlaapglYalI-----Al 444

Db 1504 GCTTTCGCCGAGATCAATGAGCCACCTGCAGCATTGAGAGGAGAAAGAACTGAAGC 1563

Qy 444 aSerAlaProThrgluIngluIngluValYalYalYabArgrThrgIn----- 460

Db 1564 CAGCACCCCACTGCTTACAGAGTGAAGCCCTTGGGGAGAGCTGAAGGCCCTGTGTCA 1623

Qy 461 ----GInserLYsProglInProglInProSerThrGlYserleuthr-GlySer----- 477

Db 1624 GGGGACACGCGGGCCCGAGAGAAACAGTAAATGTTGTTCATAATGTCCTCCGTT 1683

Qy 478 -----ProHleuSerGlueuSerValaIaenSerGlInglYalYalProa 494

Db 1684 GCTCCTTGACCAAGGAGATCTCTTGTGTTGCTCCTCCGATCATCTTACCG- 1742

Qy 494 laaenValThleuSer-----ProaInuSerp 504

Db 1743 --AGCTAACCTTCATTCCTCTTCCGTGATATCCGCCAGACCCCGAGTTGAACA 1800

Qy 504 roaepThr-----LYsGlInlaSerp 511

Db 1801 ATTCACTATCAATACTTGTGCCCTCAGGCGATGTTGCCGCAAAATCCGCTCAGT 1866

Qy 511 roleuIleSerProleuIeuaInaPglAlaIa-----CybProArgrThrArpArglua 529

Db 1861 GGTGAGCTGCTCATTTGACACCTGAGAGGATGACTCCTCCAAAACTAAGCCAGGTGG 1920

Qy 529 epGlugLYArArYgluArgrPheProThrArpLYalTYrPhelAlaYsGluals 549

Db 1921 ACCAAGGAACCCGGTATCCCATTCACCAATGGGACCCAT-----C 1962

Qy 549 erThrThgluArgrThrYrleuLYsArpLeuGlualIleThSerTrpPheglInserT 569

Db 1963 GCGGAACCATGGCAATATACCAAGTCTCTCTCATGATCA-----AAATCCA 2013

Qy 569 hrValSerLYgluArpAlaMetProgluAlaLeuLYSerleuIlePheProaInPheg 589

Db 2014 CTGACGCTTAGAGAGG-TGTTTCCCAAGAAAGAAAGGATAGGCTATGCTGTCTAA 2072

Qy 589 luProluHleYs-----PheHlThrArpInPheleuLYglu 601

Db 2073 ACAAACTGGGAAACTATTTTCTTCAGAAAGTTATTTCAAGAA 2116

RESULT 9
US-09-848-294-1
Sequence 1, Application US/09848294
Patent No. 6479640
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas K.
TITLE OF INVENTION: Isolation of A cDNA Encoding A No. 6479640el
TITLE OF INVENTION: Protein Tyrosine Phosphatase Which Localizes to Focal
FILE REFERENCE: CSHL90-04PZA
CURRENT APPLICATION NUMBER: US/09/848-294

	/	CURRENT FILING DATE:	2001-05-03	
	/	PRIOR APPLICATION NUMBER:	09/235,251	
	/	PRIOR FILING DATE:	1999-01-22	
	/	PRIOR APPLICATION NUMBER:	08/759,536	
	/	PRIOR FILING DATE:	1996-12-04	
	/	PRIOR APPLICATION NUMBER:	08/107,420	
	/	PRIOR FILING DATE:	1993-08-16	
	/	PRIOR APPLICATION NUMBER:	07/663,579	
	/	PRIOR FILING DATE:	1991-03-01	
	/	PRIOR APPLICATION NUMBER:	07/494,036	
	/	PRIOR FILING DATE:	1990-03-14	
	/	NUMBER OF SEQ ID NOS:	13	
	/	SOFTWARE:	PatSeq for Windows Version 3.0	
	/	SEQ ID NO 1		
	/	LENGTH:	3984	
	/	TYPED:	DNA	
	/	ORGANISM:	Homo sapiens	
	/	FEATURES:		
	/	NAME/KEY:	CDS	
	/	LOCATION:	(24) ... (2765)	
	/	US-09-848-294-1		
		Alignment Scores:		
		Pred. No.:	1,29e+40	Length: 3984
		Score:	541.00	Matches: 229
		Percent Similarity:	40.1%	Conservative: 151
		Best Local Similarity:	24.2%	Mismatches: 358
		Query Match:	9.9%	Indels: 210
		DB:	3	Gaps: 35
QY	US-09-555-342B-2 (1-1045) x US-09-848-294-1 (1-3984)			
QY	27 ArgGlyInLysProPheProThrProSerGlyLeuValSerIleGlnMet	46		
Db	69 CGCACCTCGGAGTTACCCAAAGAAGAAACTGCATCAGAAGTATTGGAGATCACTTT	120		
QY	47 LeuAspAspThrGlnGluAlaPheGluValProGlnArgAlaProGlyLysValLeuLeu	66		
Db	129 TTGATGTGCGTGTAAGACGACTTTAAAGTTACTTAAACAAGACACTGCCAAGTTCCTTG	188		
QY	67 AspAlaValLysAsnHisIleAsnLeuValGluGlyAspTrpPheGlyLeuGluPhePro	86		
Db	189 GATAATGTGCACAAACCACTGGGTGTGACTGAAGAAAAGAAATATTTTGTTTACAGCATGAT	246		
QY	87 AspiIeLysLysIleThrVal---TriLeuAspLeuLeuLysProIleValLysGlnIle	105		
Db	249 GACGACCTCCGTGACCTCTCTAGATGCTGGAAGAACAAACCATTAGAGAACAGTTA	308		
QY	106 ArgArgProLysHisValValValLysPheValValLysPhePheProProAspHisThr	125		
Db	309 AAAAGAGGTTTTCCCTGTAACCTGCATTTTCAGTAGAATTTTATTAAGTATCCGATCCCAAC	368		
QY	126 GlnLeuGlnGlnLysLeuThrArgTrpLeuPheAlaLeuGlnValLysGlnAspLeuAla	144		
Db	369 ACACTGCAGCAAGAAACAAACCAAGGCACTTGTTTCTTAACAACGTGAAGATGGAATTTGC	422		
QY	146 GlnGlyArgLeuThrCysAsnAspThrThrsSerAlaIleLeuLeuIleSerHisIleValGln	165		
Db	429 GAAGAGAGGTTTAACTCGCCCTCTTAAGTACGAGAGGTTTTCAGCTCTTAAGCCGTTACA	488		
QY	166 SerGlnIleGlyAspPheAspGluAlaLeuAspArgGlu--HisIleAlaLysAsnLys	184		
Db	489 TCTCATTTTGGAGACTATATATCTTCATACATCATCCAGGCTATCTTCGATATGAC	548		
QY	185 TyrIleProGlnGlnAspAlaLeuGlnAsp-----LysIleValGlnPheHis	203		
Db	549 TTTATATCCCATAAAT-----GAGGACTTTTAAACAAAGTCGAATCTCTGCAATGAG	602		
QY	202 AsnHisIleGlyGlnThrProAlaGlnSerAspPheGlnLeuLeuGlnIleAlaArgArg	221		
Db	603 CAGCAGCAGTGGGCTTAAACATCAGAGAGCAATCTCTATATCAATGACGCGGAGAC	662		
QY	222 LeuGlnMetTyrGlyIleArgLeuHisProAlaLysAspArgGlnGlyThrLysIleAsn	241		

Db	663	CTCGACTTCTAAGGATGAACTGCACAGTGTAGAGATCTTGCACAATTTAGACCTTAATG	722
Oy	242	LeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrGlyIleAsnAlaPhe	261
Db	723	ATTGGAATTGCTTCGCGGGGTCTCTGTGTACCGAAAAATACATTATGGACAAGTTTCAT	782
Oy	262	AsnThrAlaValAlaArgLeuSerPheLeuValGlyValArgPheLeuIleLeuLeuArg	281
Db	783	CCTTGGGTGAACATCTCTCAAAATTTCTTTCAAAAGAAAGTTCTTCATCATCAGCGA	842
Oy	282	ProAspAlaAsnSerAlaThrGlnAspThrLeuGluPheLeuMetAlaSerArgAspPhe	301
Db	843	CAGAAACAGGCTGAATCCAGGGAAACATATTGTGGCTTCAACATGCTGAATTACCGATCT	902
Oy	302	CysIysSerPheThrPylIleCysValGlnHisAlaPhePheArgLeuPheGluGlu	321
Db	903	TGCAGAAACTGTGTGAAAATCTGTGTGTAGCACCATACGTTCTTTACAGCA---AAGAG	959
Oy	322	ProIlyPProIlyPProIlyPProValLeu-----PheSerArgGlyIysSerPheArg	338
Db	960	CTACTACTCCAGAGAAAGATGTTCTGTCTCAGTACTGGACTATGAGGCTCTCGG-----	1013
Oy	339	PheSerGlyArgThrGlnIlyGlnValLeuAspTrValIlyGluGlyIleIlyLeuIys	358
Db	1014	-----AACACCAAAAGTCGGTAATATACCA-ATATTGCAGAAAAGGATTTGGGG	1063
Oy	359	ValGlnPheGluArgIlyHisSerIylIleHisSerIleArgSerLeuAlaSerGlnPr	378
Db	1064	GATGTTGTGGAACCCAGCCATGCGGAGATTCCTTATCAGTGAGCACTTAGAAACCAAGAG	1123
Oy	378	ThrGluLeuAsnSerGluValLeuGluGlnSerGlnSerThrSerIleThrPheGlu	398
Db	1124	TCTGCTTCTGTGTCCTCCCTCCATTATCTCCAACTGGGAGAGTCTCGAGTC-----	1175
Oy	398	YglGluIylAglUserProGlyIlyGlnSerCysArgArgGlyIlyGluPProIlyValSe	418
Db	1175	-----	1175
Oy	418	AlaGlyIylGluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnIly	438
Db	1176	-----CGGCAGAAATCCGAAGCCAGCACTCT-----	1205
Oy	438	AsnIlaAspGlyAlaAlaSerAlaProThr-----GluGluGlnGluGluValIly	456
Db	1206	-TCTGCATATACCTTGCAATGAATGACCTACATCAGGAAACGAAGATGATTTTA	1264
Oy	456	AspArgThrGlnGlnSerIlyPProGlnProProGlnProSerThrGlySerLeuThrGlu	476
Db	1265	CACGTACAGGGCTTCTGCGCCCTCA-----GACAGCATTTCTGAAGTTTCTCAGAACCG	1321
Oy	476	YSerProHisLeuSerGluLeuSerValAsnSerGlnIlyGlyAlaAlaProAlaAsnVa	496
Db	1322	AAGCCCGACCAAGAGCTTATTCGAGAACAT-----CCGGCACAAAG	1366
Oy	496	IThrLeuSerProAsnLeuSerProAspThrIlyGlnAlaSerProLeuIleSerProLe	516
Db	1367	CTACTCT-----ACCCAGAAAGTCATCCAGTCTGTCTCCATC	1405
Oy	516	WLeuAsnAspGlnAlaCysProArgThrAspAspGluAspGluArgGlyValArgPh	536
Db	1406	TTCAATGCTCCAGGCTCTGCTCACTGACGCGCTTGATCAGCAGCTC-----	1454
Oy	536	eProThrArgPylAlaIlyrPheIleAlaIlyGluValSerThrThrGluIlyThrIlyLe	556
Db	1455	---TTAAGATGACTTCCACAGGATACCAAGGGGGCTCCACCGAGCGCCAG-----	1505
Oy	556	UlyAspArgGluValIleThrSerTrpPheGlnSerThrValSerIlyGluAspAlaMe	576
Db	1506	-----CAGTACTACTGTGCAAGATGATGAATGG	1534
Oy	576	tProGluAlaLeuIlySerLeu---IlePheProAsnPheGluPProLeuHisIlyrPheHis	595

Db	1535	TGACAGCTACTTAAGTCTTGAATCCGTATCCATGACACCCAGATGAAGAT-----GGAAATTTGGC	1588
Qy	595	ethrAspPheLeuValIeGlnGlnArgLeuAlaLeuTrpGluValArgSerAsnAl	615
Db	1589	ATTATATCTTAAGGAGAGAGATGTGATCAAAAGATGCTCTTGTGTGATCAAGATAAACCC	1648
Qy	615	agIn-----IleArgSerTrpGlnArgIleGluAspVala	626
Db	1649	AGAGTCACTGCGGACACTGCACTTCCATAGCTGAAGAGAGGATCAAAATCGTGTAT	1708
Qy	626	I MetLeuValAsnIleGlnGlnMetCysHis-----LeuAlaIAsnI	640
Db	1709	CAATGCGCGGACACTTCAAGAACACACAGCATGACCAAGTGTATGTTCATCAAAAGCCAG	1768
Qy	640	sLeuTrpLysHisSerGlnAlaLeuGlnAlaLeuGlnAlaValIleLysSerSerArgAr	660
Db	1769	CGGGAGTCCCACTCAAGGAGCTG---GCCCTGTATGACGAGGAGAGCTGTCCGCTC	1825
Qy	660	gLeuGlnAsnPheCysArgArgPheGlnLeuGlnValCysTrpTrpLeuProLeuAsnTrh	680
Db	1826	ATTTCCTCACTTCAAGTCTGAATGAATGAATGAATCAAGCTT-----	1865
Qy	680	rPheLeuLeuArgProLeuHisArgLeuMetHisTrpTrpGlnValLeuGlnValArgLeuCy	700
Db	1866	-----TTCGCCAGAACCTTTTCCCATGTG	1891
Qy	700	sLysHisAspProSerSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGlnI	720
Db	1892	T-----CGGAGGGGTGGGGAC-----	1907
Qy	720	eThrGlnMetValAlaGlnLeuHisGlyTrpMetIleLysMetGlu-----	735
Db	1908	-----ACTTTGGAGGGATCCATGACAGCTAAAGAAAGGCGCTCGAAAG	1951
Qy	736	-----AsnPheGlnLysLeuHisGlnLeuLysValLysValAspLeuIleGlyI	750
Db	1952	CGGAGCGTGTGATCCAGTTTGGACACTTCAAGAAAGAAAG-----	1994
Qy	750	eAspAsnLeuValAlaProGlyArgGlnPheIleArgLeuGlnLysLeuSerLeuLysLeuSe	770
Db	1995	-----CCAGTTTGGGCCATC-----ACGTTTGAAGAGCTGCC	2026
Qy	770	gGlyLysGlnLeuGlnGlnArgMetPhePheLeuPheAsnAspValLeuLeuTrp--Th	789
Db	2027	TCAAAATTTGGACAAAAACGA-----TATTAAGATGTGCTGCTTATGACAC	2074
Qy	789	rSerArgGlyLeuThrAlaSerAsnGlnPheLysValHisGly-----G	804
Db	2075	CACCCGGGTATTAATGACAGGAAATGAAGATTATTAATGACAACTTACCTGAACATGACA	2134
Qy	804	nLeuProLeuTrpGlnMetThrIleGlnGlnLysSerGlnAspGlnTrpGlnValProHis--	823
Db	2135	AATTCCTGCTCTAACCTTGGAACAAAGTACATGCGCACTGAGGGGCCCTCGCCGATAC	2194
Qy	824	CysLeu-----ThrLeuArgGlnArgGlnSerIleIleValAlaAlaase	839
Db	2195	CTGTGCAAGTTTGGCAGGTGTGCTGGAGATCAAGAATGTGCATCTCATTTGC-----	2246
Qy	839	rSerArgSerGlnMetGlnLysTrpValGlnAspIleGlnMetKAlaIleAspLeuAlaG	859
Db	2247	-----ATGTTGACGACTTCACAGA	2266
Qy	859	sLysSerSerProAlaProGlnPheLeuAlaSerSerProProAsp-----	875
Db	2267	ACGAGGGCGGACCAATGTCAACCACTGTG---CAAGATCCCCCGACGTCATGAACCA	2323
Qy	876	-----AsnLysSerProAspGlnAlaThrAlaAlaAspGlnLysG	890
Db	2324	CGGCGGCTTTCACATCCAGTGTCAGTGACAGAGACTGACACCATGCTATGTGCTCCGAGA	2383
Qy	890	uAspAspLeuSerAlaSerArgThrSerLeuGlnValGlnAlaProHisArgGlyAsnTrh	910
Db	2384	AATGTGTGTCAAAACCCCAAGACCGGGGAGAACACACAGTGCACAT-----	2432

```

Qy      910 rmevalhisvalcysrttr 916
Db      2433 -ctccagtaactgctcagatgg 2450

RESULT 10
US-09-949-016-3593
; Sequence 3593, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3593
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3593

Alignment Scores:
Score: 6.28e-38      Length: 2156
Percent Similarity: 508.00      Matches: 148
Best Local Similarity: 47.0%      Conservative: 78
Query Match: 9.3%      Mismatches: 198
                        Indels: 57
                        Gaps: 14

US-09-555-342B-2 (1-1045) x US-09-949-016-3593 (1-2156)
Qy      27 ArgGlyGlnIysProProThrProSerGlyLysLeuValSerIleValIleGlnMet 46
Db      63 CGCACCTCGGAGTACCAGAAAGAGAAAGAACTGATCAGAGTCATTGGAGATCCACTTT 122
Qy      47 LeuAspThrGlnGlnValPheGlnValProGlnArgAlaProGlyLysValLeuLeu 66
Db      123 TTGATGCGCGTGTACAGACCTTTAAAGTTACTAAACAGACCTGCGCAGGTTCTTCTG 182
Qy      67 AspAlaValCysAsnHisIleuAsnLeuValGlnIysPheGlyLeuGlnPhePro 86
Db      183 GATATGGTGACAAACCACTGGGTGTGATGTAAGAAAGATATTTGGTTTACAGCATGAT 242
Qy      87 AspHisIleValIleThrVal---TrpLeuAspLeuLeuLysProIleValIleGlnIle 105
Db      243 GAGGACCTCGGAGACTCTCTCAGATGGCTGGAAGCAAGCAATCAGGAAACAGATTA 302
Qy      106 ArgArgProLysHisValValIleValPheValValIlyPhePheProProAspHisThr 125
Db      303 AAAGGAGGTTTCCCTGTACCTCATTTTTCGATAGATTTTAACTGATCCCAAC 362
Qy      126 GlnLeuGlnGlnIleuLeuThrArgTrpLeuPheAlaLeuGlnValIlyGlnAspLeuAla 145
Db      363 ACATGCGAGCAAGCAACCAAGGCACTGTATTTCTTACAACTGAAAGATGATATTTGC 422
Qy      146 GlnGlyArgLeuThrCysAsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGln 165
Db      423 GAAGGAGGTTTAACTGCGCTCTTAACCTCAGACAGGAGTTTAAAGTCCGATAGCCGTA 482
Qy      166 SerGlnIleGlyAspPheAspGlnAlaLeuAspArgGlu---HisLeuAlaLysAsnLys 184
Db      483 TCTCATTTTGGAGACTTAATTTCTTCATATCATTCACAGGCTATCTTTCCGATAGTCAC 542
Qy      185 TyrIleProGlnGlnAspAlaLeuGlnAsp-----LysIleValGlnPheHisHis 201

```

```

Db      543 TTTATACCCGATCAAAAT-----GAGGACTTTTAAACAAAGTCAATCTGTCATGAG 596
Qy      202 AsnHisIleGlyGlnThrProAlaGlnSerAspPheGlnLeuGlnIleAlaArg 221
Db      597 CAGCAGTGGCGGTAAACAACTAGAGCAAGAACTCTGTATATCAATAGCCGCGAGCC 656
Qy      222 LeuGlnMetTyrGlyIleArgLeuHisProAlaLysAspArgGlnGlyThrIlyIleAsn 241
Db      657 CTCGACTTCTATGAGAGTAAGAACTGACAGTGGTAGGATCTGCAAAATTTAGACTATAG 716
Qy      242 LeuAlaValAlaAsnThrGlyIleLeuValPheGlnIlyPheThrIlyIleAsnAlaPhe 261
Db      717 ATTGGAATTCCTCCGGGGGTGCTGCTGTGACCGAAATACATTTGCACAAAGTTTCTAT 776
Qy      262 AsnTrpAlaLysValArgLysLeuSerPheLysArgLysArgPheLysIleLysLeuArg 281
Db      777 CTTGGGTGAACAATCTCAAAATTTCTTCAAAAGGAAAAAGTTCTTCATCATCAGCGA 836
Qy      282 ProAspAlaAsnSerAlaTyrGlnAspThrLeuGlnPheLeuMetAlaSerArgAspPhe 301
Db      837 CAGAAACAGCTGAATCAGGGAACATATTTGTGCTTCAACATGCTGAATTCAGATCT 896
Qy      302 CysLysSerPheTrpLysIleCysValGlnHisIleAlaPhePheArgLeuPheGlnGlu 321
Db      897 TGCAAAACCTTGCGAAATCTCGTGTGACACCATACGTTCTTTCAGGCA---AAGAG 953
Qy      322 ProLysProLysProLysProValLeu-----PheSerArgLysSerPheArg 338
Db      954 CTACTACCTCAGGAAAGAAATGTTCTGTCTCAGTACGACTATGGGCTCTCGG----- 1007
Qy      339 PheSerIlyArgThrGlnLysGlnValLeuAsp---TyrValLysGlu-----GlyGly 355
Db      1008 -----AACCACAAAAGTCGTAATATACCAATATTCGAAAAGGTGATGGGCGG 1058
Qy      356 -----HisLysValGlnPheGlnArgLysHisSerLysIle 368
Db      1059 ATGCTGTGAACCCAGCCATCGGAGATCCTTATCAGTGGAGCACTTAGAAACCAAGAGT 1118
Qy      369 HisSerIleArgSer-LeuAlaSerGlnProThrGlnLeuAsnSerGlnValLeuGln 388
Db      1119 CTGCTCTCTGTTCCCTCCCATTAATCTCCCACT----- 1152
Qy      388 nSerGlnGlnSerThrSerLeuThrPheGlyGlnGlyAlaGlnSerProGlyGlnIle 408
Db      1153 -----GGGAGGAGTGGATCAAAAGATGCCTC 1178
Qy      408 rCysArgArgGlyLysGlnProLysValSerAlaGlyLysProGlySerHisPProSerPr 428
Db      1179 TTGTGTATCAAGATTAACCCAGATCAGCTCGGACAC-----CTGCAT---TCTTA 1229
Qy      428 AlaProArgArgSerProAlaGlnAsnLysGlnAlaAspGlyAlaAlaSerAlaProTh 448
Db      1230 AGCTGAACGAAGGGGATCAATCGTGTATTAATGACCGGACATCTCAGAACACAGC 1289
Qy      448 rGlnGlnGlnGlnValValIlyLysAspArgThrGlnGlnSerLysPheGlnProGlnPro 468
Db      1290 ATGACCAAGTGTGATGTT-----CATTAAGCCAGCCGGGAGT 1328
Qy      468 nProSerThrGlySerLeuThrGlySerProHisLeuSerGlnLeuSerValAsnSerG 488
Db      1329 CCCACTCAGCGAGCTGCGCTGTATATCAGAGGAAAGCACTCTACAAAAAAAGCCA 1388
Qy      488 n 488
Db      1389 G 1389

RESULT 11
US-09-270-767-1995
; Sequence 1995, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.

```

```

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1995
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; US-09-270-767-1995

Alignment Scores:
Pred. No.: 1.67e-38 Length: 576
Score: 503.50 Matches: 92
Percent Similarity: 73.8% Conservative: 32
Best Local Similarity: 54.8% Mismatches: 43
Query Match: 9.2% Indels: 1
DB: Gaps: 1

US-09-555-342B-2 (1-1045) x US-09-270-767-1995 (1-576)

QY 200 HIEHIAASHNHSIIEGLYGINThrProAlaGluSerAspPheGlnLeuLeuGluIleAla 219
DB 6 CACTAAAGCATGTGTGTCATCCCGCGAGCGGACCTTAACCTCTTGAGACGGCT 65
QY 220 ARGARGLeuGlnMetTYrGlyIleArgLeuHISProAlaIleuValAspArgGluIYThrLys 239
DB 66 CGAGGTGTGAGCTGTATGGCATGAATAATGCATCCGCGAAGGTGTGAAAGGGGTGCCG 125
QY 240 ILEASNLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsn 259
DB 126 CTTAATCTGGCTGTGGCCACATGGGCATCACAGTCTTCCAGAAACATCAGCGGATCAC 185
QY 260 ALPheAsnTrpAlaLysValArgLysLeuSerPheLysArgLysArgPheLeuIleLys 279
DB 186 ACCTTCTCGGGGTGAAGATACGCAAGATTTCTTCAAGCGCAAGCATTTCTGGTCAAA 245
QY 280 LeuArgProAspAlaAsnSerAlaTYrGlnAspThrLeuGlnPheLeuMetAlaSerArg 299
DB 246 CTCGATCCGAGGATATGATATTACAAAGATACCGGATTTCTTCCAGAGGTCCG 305
QY 300 PheSerGlyArgThrGlnLysGlnValLeuAspTYrValLysGluIYglnHISLysLys 358
DB 426 TATAGCGGAAAAACCCAGAAAGCATTTATCGAGTTCTTCCGAAATAATTATGGAAGCGC 485
QY 359 ValGlnPheGlnArgLysHisSer 366
DB 486 CAAACTTCCAAAGGTCTCAGTCA 509

RESULT 12
US-09-270-767-17277
; Sequence 17277, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17277
; LENGTH: 576
; TYPE: DNA
```

```

; ORGANISM: Drosophila melanogaster
; US-09-270-767-17277

Alignment Scores:
Pred. No.: 1.67e-38 Length: 576
Score: 503.50 Matches: 92
Percent Similarity: 73.8% Conservative: 32
Best Local Similarity: 54.8% Mismatches: 43
Query Match: 9.2% Indels: 1
DB: Gaps: 1

US-09-555-342B-2 (1-1045) x US-09-270-767-17277 (1-576)

QY 200 HIEHIAASHNHSIIEGLYGINThrProAlaGluSerAspPheGlnLeuLeuGluIleAla 219
DB 6 CACTAAAGCATGTGTGTCATCCCGCGAGCGGACCTTAACCTCTTGAGACGGCT 65
QY 220 ARGARGLeuGlnMetTYrGlyIleArgLeuHISProAlaIleuValAspArgGluIYThrLys 239
DB 66 CGAGGTGTGAGCTGTATGGCATGAATAATGCATCCGCGAAGGTGTGAAAGGGGTGCCG 125
QY 240 ILEASNLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsn 259
DB 126 CTTAATCTGGCTGTGGCCACATGGGCATCACAGTCTTCCAGAAACATCAGCGGATCAC 185
QY 260 ALPheAsnTrpAlaLysValArgLysLeuSerPheLysArgLysArgPheLeuIleLys 279
DB 186 ACCTTCTCGGGGTGAAGATACGCAAGATTTCTTCAAGCGCAAGCATTTCTGGTCAAA 245
QY 280 LeuArgProAspAlaAsnSerAlaTYrGlnAspThrLeuGlnPheLeuMetAlaSerArg 299
DB 246 CTCGATCCGAGGATATGATATTACAAAGATACCGGATTTCTTCCAGAGGTCCG 305
QY 300 PheSerGlyArgThrGlnLysGlnValLeuAspTYrValLysGluIYglnHISLysLys 358
DB 426 TATAGCGGAAAAACCCAGAAAGCATTTATCGAGTTCTTCCGAAATAATTATGGAAGCGC 485
QY 359 ValGlnPheGlnArgLysHisSer 366
DB 486 CAAACTTCCAAAGGTCTCAGTCA 509

RESULT 13
US-10-094-749-1401
; Sequence 1401, Application US/10094749
; Patent No. 6979557
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
```


FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIORITY APPLICATION NUMBER: 60/350,435
PRIORITY FILING DATE: 2002-01-24
PRIORITY APPLICATION NUMBER: JP 2001-328381
PRIORITY FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1401
LENGTH: 2931
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-1401

Alignment Scores:
Pred. No.: 1,09e-36 Length: 2931
Score: 497.50 Matches: 207
Percent Similarity: 41.3% Conservative: 161
Best Local Similarity: 23.2% Mismatches: 377
Query Match: 9.1% Indels: 146
DB: 4 Gaps: 31

US-09-555-342B-2 (1-1045) x US-10-094-749-1401 (1-2931)

QY 248 G1y1leuval1pneinglypneThry1leahna1pnean----- 262
DB 85 GGGCTGGCTGGCTCGCTGGAGAAACAATGAGAGGATTTACTTCAACCATTTGTGAC 144
QY 263 ---Trr1alalyVal1ArglyLeuSerPheLybArghyPheLeu1leLyLeuArg 281
DB 145 TTCTGGCCAGAGAAATTAATTCAGCTTTTAAGGCGAGTACCTGTCCAAAGATGCG 204
QY 282 ProAer1aenSer1aTyrg1naBrThreug1upheLeuMet1aSer----- 298
DB 205 -----TTTAACTTCCACCTTGTCCACAGACAGACACACACACTGTGT 246
QY 299 -----ArgAerPheCylybSerPheTrpLy1leCy1ValG1uH1 312
DB 247 TGGTGAATGTATCTGAAGAGAGCTCAGGAGTAATGGGAACGGCCAGCAAAACA 306
QY 312 vH1a1A1PhePheArglyPheug1upProLyProLy-----ProLyPProVal1e 330
DB 307 CTGACGT-----GCAAGTCCAAACCCCAAGTGGCTCCAAAGCCATTACA 351
QY 330 upHe-----SerArgLySerSerPheArgPheSerG1yArg-----Th 343
DB 352 CCTGCAAAATTCACCTTCGTCATATACACCAAAACCCCGAGATAAAGCTTACCTAG 411
QY 343 rg1nLyug1nVal1euarTyrg1valLyvg1uLyg1nH1bLybVal1G1nPhug1uAr 363
DB 412 TGCAAAACCAAGATGAGAGAAATTAACCTGCTGCTTGTGTGTCMAAAGAA-- 469
QY 363 gLyvH1eSerLy1leH1eSer1leArGserLeu1aSerG1nProThrg1uLeuAnse 383
DB 470 -AAACCAAGTAAATGATCAATCATCAAGTCCGTTTGAAGAGAGCAAGCTTATTAACA 528
QY 383 rg1nVal1eug1nSerG1nSerThSerLeuThrPheG1uLyg1u1aG1uSe 403
DB 529 TTTAAGTATTTGAAGAAAGAGTCTGCTGTAACCTA-----AATGCTCTTAAGAC 579
QY 403 rPrroG1yG1nSerCybArghyG1yLyvg1uProLy-----ValSer1aG1yG1uPr 422
DB 580 CCCAGAGAGGATGATGACCAACCAACCTCAACAAATCTCTCCCGACACTTGGC 639
QY 422 oG1ySerH1eProSerPro1aProArghySerPro1aG1yAnLyvg1n1aAerG1 442
DB 640 ACAAGGCGAGAAATGATGCAATTAAGATCAAGGTCGACAGACTTGTGTGCGCAACGG 699
QY 442 y1a1a1Ser1aProThrg1uG1uG1uG1uG1uVal1yAerArghyThrg1nG1nSe 462
DB 700 TGTAAATGCGACAAACCAAGATGATGAGAGAGAAAGTCCCACTTAACTGC 759

QY 462 rLyvProG1nProProG1nProSerThrg1ySerLeuThG1ySerProH1eLeuSerG1 482
DB 760 A-----GATACCTTCATTAAGCTTTCGAACCTTCCTGTGATAC 798
QY 482 uLeuSerValAsn-----SerG1nLyg1Val1aPro1aAanValThreLeuSe 499
DB 799 GCACATAGTGAATGAGAAAGATGAATGCAAGCTTCCTGCATCACCACCAACAGAGA 858
QY 499 rPro-----AsnLeuSerProAerThryvg1nH1aSerProLeu1leSerProLe 516
DB 859 TAGCTGTAGTGAATTCCTCT--GACAGTACGTACAGGATTCACGATTAAGCCAGT 915
QY 516 uLeu-----AsnAerG1n1aCyvProArghThrAerArgh1uSerG1uG1 531
DB 916 GCTCCCTTGAAGAGAGAGGCGAGAAACAGAAACCAAGTACAGAGGAGAAATGG 975
QY 531 Y-----ArgArghyArghPheProth 538
DB 976 GAAAGCCCTGTGAATCGAGACGTGACACAGACCATGAGATGAAGAGACTAATGA 1035
QY 538 rAerLyA1aTyrg1he1le1aLyvg1uValSerThThrg1uArghThryrLeuLyAs 558
DB 1036 GCAAAATTCACAAATTAAGCATGAACTTTGCTTACGAAGACTTAATGCAACCG 1095
QY 558 rLeuG1uVal1leH1eSerTrpPheG1nSerThrValSerLyvg1u-----AerA1 575
DB 1096 ACTGACCTCTTGATAGATGATTTATTTGCAACCTTGGAGAAACAAACGAGCTC 1155
QY 575 aMerProG1u1a1eLybSerLeu1lePheProAerPheG1uProLeuH1yLyvPheH1 595
DB 1156 GTTTCACGACAGATGATGTAATAAATCTTTCCAAATTTCAATCAATTAATGCTTCCA 1215
QY 595 vThrAerPhe---LeuLyvg1u1leG1uG1nArghLeu1aLeuTrG1uG1uArghSerAr 614
DB 1216 TAGTAAATTCCTTGTCCAGAGCTGAGAAACGATCAAGATTAAGGGA----- 1264
QY 614 n1aG1n1leArghArghyrg1nArgh1leG1yAerValMetLeuLybAn1leG1nG1yMe 634
DB 1265 -----ACTACCTTCGAATTTGAAGATCTTCGAGAAATTTGGACCACTTCT 1311
QY 634 tLyvH1eLeu1a1a1aH1eSerTrpLyvH1eSerG1u1a1eug1u1a1eug1uAnG1 654
DB 1312 TAAGATGTAGAGAAATGTAAGAAAGATTGTAATGCAATGCAATGATTTAAATCAT 1371
QY 654 y1leLybSerSerArghyrg1uG1uAnPheCybArghArPheG1uG1uLyvVal1Cy 674
DB 1372 GACGAAACGTATTTCCCAAGTTCAAATCATGAGTGTGAAGAAATTCAGAAACGAAATCTG 1431
QY 674 s---Tyrg1uProLeuAnThrPheLeu1euarGProLeuH1aArghLeuMetH1yTyrg1 693
DB 1432 TGGAGCTTAACCTTGAAGATCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1491
QY 693 vG1nVal1eug1uArghLeuCybLyvH1eSerProProSerH1e1a1aerPheArghCy 713
DB 1492 GAGTCTCTTAAGACATGATGAAGAAATGCTCTGATTCCTGAGTCCGAGATATGTC 1551
QY 713 vArgh1a1a1eLyb1aG1u1leThrg1uMetVal1aG1nLeuH1eG1yThre1leLy 733
DB 1552 TAAAAATTCACCTTGAATTTATCTACAGACAGACCATTTGATGTCATTAAGGAA 1611
QY 733 vMetG1uAnPheG1nLybLeuH1eG1uLeuLybLybArghLeu1leG1y1leAerPhe 753
DB 1612 AATGAGAACTTAAGAACTCTTAAGATTTTAAGATGTGAGAGAGAGAAAGACAT 1671
QY 753 vVal1a1ProG1yArgh1uPhe1leArghLeuG1ySerLeuSerLybLeuSerG1yLyv-- 772
DB 1672 TGTAAACCTTCAATGAATTAATAAGAAAGACAGATCTCAATACCTGCTGCGAA 1731
QY 773 ---G1yLeuG1nArghArghPheLeuPheAnPheVal1eLeuTyrg1---ThrSe 790
DB 1732 CACTTACGACAAAGAGCTTACCTTTTATTCACCAATCATGTGTGTGATCGTGTGCC 1791
QY 790 rArghLyLeuThAlaSerAnG1nPheLybVal1n1eG1yG1nLeuProLeuTyrg1yMe 810

Db 1792 CAAATTACGCTTGGTAGGCTCTAAATTTCACAGTTCCAGACCAGGCTTGGCATTTGGAAT 1851
 Oy 810 tThrIleGluIuSerGIuAspGIuTrpGIuValProHiScyLeuThrLeuArgIyGI 830
 Db 1852 GAAATTGTGAGACTCAAAATGAA---GAAATTCACATACTTTCAGAGTGTCTGGGAA 1908
 Oy 830 nArgInserIleIleValAlaIleSerSerxArgSerGIuMetGIuIlyTrpValGIuAs 850
 Db 1909 AGAGAACACCTCGAACTTCAGGCCAGTTCCTCGCAGACGAAAGAAAGAAATGGATCAAGGC 1968
 Oy 850 pIleGluMetAlaIleAspLeuAlaGIuLySerSerSerProAlaProGIuPheLeuAl 870
 Db 1969 CCTTCAGAAACCATCGATCGATCTTTTCATCAAAAGGACGAAACCTTCAGAAATGCATTTGC 2028
 Oy 870 aSerSerProProAlaPheLeuLySerProAspGIuAlaThrAlaIleAsp----- 886
 Db 2029 AAAG-----GATTAATGACATTCACCTCAGAGGTTTCTACTCTGACGCTAGGAGAAAG 2079
 Oy 887 -----GInGIuSerGIuAspLeu-----SerAlaSerx 897
 Db 2080 AGCCCCAGATGATCCGAGATATATACATGACACATGTGTATGAAATGTAAAGAACCTTT 2139
 Oy 897 gThrSerLeuGIuArgGIuAlaProHiSaArgIyAsnThrMetValIleValCysTrpHi 917
 Db 2140 CAATCAGCTACGACACAGAGGAGCATCATTTGTGCAGCATGTGATATATGTGTGTGGAA 2199
 Oy 917 SaArgAsnThrSerValSerMetValAspHeser----- 928
 Db 2200 ATGCTCCGACTACAAAGCTCAACTGAATATGATGTGTAAATTGACGAAAGTTGTAA 2259
 Oy 928 ----- 928
 Db 2260 AGACTGTTATCAATATCATTAATGTGATTACACAGACTGAAAGAAAGAAAGAAAGAAAT 2319
 Oy 929 -lIleAlaValGIuAsn---GInLeuSerGIuAsn-----LeuLeuArgIy 942
 Db 2320 TTTAAGATGTGAATCAGACGAAAGTATCTTGAAACAGTGTGTGTGACAGCTTTCTTCAGTA 2379
 Oy 942 aPheLyAsnSerAsnGIuTrpGIuLySerLeuTrpValIlePheThrAsn-----Ph 959
 Db 2380 TATGAGAAAGTCAAAACCTTGGCAGAAACCTTGTTGTGTATCCCAACAGAACCCCTCT 2439
 Oy 959 eCysLeuPhePheTrpLySerSerHiGIuAspAsnHiSProLeuAlaSerLeuProLeu 979
 Db 2440 TGTGCTGTACATGTATGTGTGCCCCCCAGACGTCAGACAGCCACGACCATTCACCTTCT 2499
 Oy 979 uGIuTrpSerLeuThrIleProSerGIuSerGIuAsnIleGInLyAspTrpValPheLy 999
 Db 2500 GGGCTATGTGTGTGATGAATGCCAGAGAGCCGACGACTG-----CCACACAGTTTCAA 2553
 Oy 999 aLeuHiAspLeuLySerHiAsnValTrpTrpPheArgAlaGIuSerGIuTrpThrPheGIuAr 1019
 Db 2554 ACTGACCCGAGTCAATGTCGTGACACAGCTTTGTCTGCAGACAGTGAAGAACTGAAGCAGAA 2613
 Oy 1019 gTrpMetGIuValIleArgSerAlaThrSer 1029
 Db 2614 GTGGCTGAAGTATCTTTTAGCTGTCA 2644
 RESULT 14
 ; US-09-949-016-5747
 ; Sequence 5747, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CLO001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768

```

; PRIOR FILLING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5747
; LENGTH: 4272
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5747

Alignment Scores:
Pred. No.:      3,44e-31      Length:      4272
Score:          443.00       Matches:     180
Percent Similarity: 39.1%    Conservative: 111
Best Local Similarity: 24.2% Mismatches:   302
Query Match:      8.1%      Indels:        153
DB:               3         Gaps:           22

US-09-555-342B-2 (1-1045) x US-09-949-016-5747 (1-4272)

Oy      401 AAlagUserProGlyginSerCybaArgLyLysglUrolysValserAalagly 420
Db      1438 GCCCAGGTCTCCCGACCAAGGCCAGTATG-----TGCCACAGC 1473

Oy      421 GlUProGlySerHisPProserProalaPro-----              430
Db      1474 CAACCTGCAGACCACAGTCGCCAACGTCCCCCGAAGGTGAAGCCCTCCGGCTGTTC 1533

Oy      431 -----ArgArgSerProAlaGlyAsnLySglnAlaasp 441
Db      1534 TGCTGGCTCTGGGCCCCGGGGA CGGTAGAAGGTGCCAACCGGACAGCGCATTTGATA 1593

Oy      442 GLYAlaIalaserialProThrGluglUGluGlUGluValValylaspaRythngIngn 461
Db      1594 GCATCACACTCGCCATCCAAACGCGAGG-----                1620

Oy      462 SerLysePrOgInProFrogInProSerThrGlySerLeuthrhrgLySerProHleUser 481
Db      1621 -----AAACCTGCTTCCTCACTAGTAAGCGGGCCCCCA----- 1653

Oy      482 GlULeuSerValanSerGingLygLYAlaIaAProAlaAnValThrlLeuSerProasn 501
Db      1654 ----GCCACAGCCTCTGCCCTGGGGCCCCCTGCCGTAGTGtgcctgttgcccttg- 1706

Oy      502 LeuSerProkasythrLySglnAlaserProleulleSerProleulemaaspGlnAla 521
Db      1707 GCCCAGCCCACCGCGCTGGCTCCCAAAGATTGACA CTACTGACGAGAGAGAC-- 1763

Oy      522 CySProArghThraPaRaRGluAaPRGLuGYARgaRgyLaRgrPhPeProTrn----- 538
Db      1764 -----GACCAAGACGACGAGAGAAAGABAAGACAGAAATCCCAAGTCCCCCTGATG 1814

Oy      539 -----AspLyahLatyrhellleaLYsgLu 547
Db      1815 GAGAGACAGAGATCTGTGAGATTGACTGTGCACAAAAAGGTGTTTCATTGACATGCCAAATGAG 1874

Oy      548 ValSerThrhrgLiuaRghTryTYrLEuLYaaRPLeugluValilleThrsertTPheGln 567
Db      1875 CTCCTGCMAAcTAGAAGGCCCTACcggtttcCAAgGCTCCATGTCCTCGAGTCAAGstgttgt 1934

Oy      568 SerThralSerLySgLn-----AspAlaMetProGluAlaLeuLySerLeu 583
Db      1935 GCCCGGCTGTGGAGAAGAACGTGGAAACCGGACAGTTCTTCCCGGCGGACGTTGTCCAGGC 1994

Oy      584 IllePheProAnpheGluProleuNileLYsrPheNiVThraenPhe---LeuLYsgLIle 602
Db      1995 ATCTTCTCTAAcATCTGECTCACTATGTCTTCACACAGAGTTCCTGCTGCCTGAAGCTA 2054

Oy      603 GluInLlrgeulaaleutrpGluGIYAArgSezaSnlaGlnlleargAaprytyGlnArg 622
Db      2055 GAGNAAGCGCATGAGGAGATGGAG-----CGCTATCCAGC 2090
```

QY	623	llleglyavpvalmleuyleuyleuamllleqnlglmetlyvnhvleuallalahlileutrp	642
Db	2091	ATTTGAGACATCTCTCAGAAATCGGCCCTTCTCAAGATGATGATGATGATGATGATG	2150
QY	643	lyvnhlseserqlualaleuqlualaleuqluamnglylleylserSerArglaleuqlu	662
Db	2151	AACTTGAACGGGGCGGTGGAGCTGGTCTCAACACCTGGACAGAGCGCTTCCACCAAGTTTAA	2210
QY	663	AsnpheCyvArArgpPheglululeuqlulnyvalCyv---TytleuProleuAhntrphe	681
Db	2211	GTCAATCATCCATGAGGTGCAGAAAGAGAAAGAGCGCTTGGCAACCTGACATTCAGACACAC	2270
QY	662	leuLeuArpProleuNhlsArgleuNehlsTyTlyvsglnvalleuqluArgleuCylyv	701
Db	2271	ATGCTTGAGCGCTGTGCAGCGCATCCCGCGCTGTGAGCTTCTTCAAGATATCTGTTA	2330
QY	702	HlsHlsrProProsenhlsalAsrPheArpCyvArGlnAlaleuAlaglnllehr	721
Db	2331	AACTGCCCCCATGCTCCCGGACAGCAAGATGCCAAAGATCTTGAGACTGATGCGC	2390
QY	722	GlumetvalalaglnleuNhlsgllyThrmetlleylsmetGluAsrPheglulnyleuhs	741
Db	2391	ACAGAGAGAGACATCGAATCTGCCATCCGCAAAATGAGAGCAATGACATAGCTGCTG	2450
QY	742	gluleuLylyvArleuLeuilegyltleAsrAneleuvalProgllyArglnlphelle	761
Db	2451	AAAGTATATGAGCTGTGAGGGGGCAGAGAGCATTTGTCAGGCCCCACCAAGAGCTCAT	2510
QY	762	ArgleuqllyserleuSerlyleuhserylyvsgly-----leuqlnqlnArpNehphe	779
Db	2511	AAAGAAAGCCACATCTTAAGCTGTGCACAAAGATGGAGACCATCAAGACCGATACCTC	2570
QY	780	PheleuPheAsrAerValleuLeuTyThrSerArglyleuThrAlaSerAsrqln---	798
Db	2571	ATACATATTCAAAGACCGCGCTCTTATCTGCGTGGCCAGAGCTGGCGCTCTTGGCCAAAG	2630
QY	799	PheLyvalNhlsGlylnleuProleuTyTlygmethrlleglnluserGlnuAsrqlu	818
Db	2631	TTTAGCGTGGGGGACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2687
QY	819	TrrgllyvalProNhlsCyaleuThrlleuArgllylnArglnglnserllelvalAlAla	838
Db	2688	CTCAATCTGCGCTCGAACCCTTCTGTGTGCAGAAAGACGCGCTCTCTGACCTCAGGCC	2747
QY	839	SerSerArpSerGluMetGlnuLeuTyTrpAlaglnuArllleglnlmevala-----	854
Db	2748	AGGATGAGAGAGAAAGAAAGACTGGGTCCAGGCCATCAACTCCACCTCTTGAAGAT	2807
QY	854	-----	854
Db	2808	GAACAGACGCTGAGACTTTCAACTGTGTAACATCAAAACAGAGAAAGATGAAGACACC	2867
QY	855	-----llehpheleuAlaglnlySerSerSerProAlaProglu--	867
Db	2868	CCACCCCACTCTCCAAAGGTGATCTTGGAAAGCGGACCATTCAGGCCATCCGGGAAAG	2927
QY	868	-----PheleuAlaserSerProPheArpAlyv---SerProArpqlualAth	883
Db	2928	GAATGCACCATGTGCATCGCTGCCAGAGCGCTTCAATCTTCAATCCAAACAGAGCAC	2987
QY	883	AlAlAlaArpqlnqlu-----serGlnuAsrPheuleuSerAlaserArqlth	898
Db	2988	CACGTCAAGAGCGCTGGGAGCATGTGGTTTGTGGAAAGTGTCCAGATTCGGGGCGCGCTC	3047
QY	898	rSerleuqluArglnAlarPonhlsArglylAsrThrmetValNhlsvalCyvTrpNhls--	917
Db	3048	GTCTATGTCAACAAACCGCTCCAAAC--CTGTGTGTGCACTGATGTTCATGTGGCTTGAC	3106
QY	918	-----ArgnAhntrSerValse	923
Db	3107	GGTGCCTTGGAGACACTCAAGCTGTGCAGACAGATCAACCCCAAGCGCGGAGACTCATCTT	3166
QY	923	rMetValAsrPheSerlleAlAlAlaGlnuAsrqlnleuhserylyvAneleuArqllyvph	943

```

Db      3167 GGAGAAAGCGCTCAGGCGCTGCAGAGAAACAGCCTCATCTGCAGCTTCCGCACTACAT 3226
        |||...||| ||||| ||| :||| :||| |||
Qy      943 e---lysanserAenGIYTPGlnLYLeuTriValAlPheThrAen-----Ph 959
        |||...||| ||||| ||| :||| :||| |||
Db      3227 GGAGAAAGGCTGGCAAAAGATGGGACACAGGACATGGTTCGTGTGCTCCCTGAAAAATGAACCTT 3266
        |||...||| ||||| ||| :||| :||| |||
Qy      959 eCysLeuPhePheTYrLYsSerHISglnAspAsthIaProLeuAlaSerLeuProLeu 979
        |||...||| ||||| ||| :||| :||| |||
Db      3287 GGTCGTGATATCTACGAGACCCCTCAGGATGTGAAACCCAGCGACCTGCCCCCTCAT 3346
        |||...||| ||||| ||| :||| :||| |||
Qy      979 uGIYrSerLeuThrIleProSerGIUsErGIUsnIleGlnLYsAspTYrValPheLY 999
        |||...||| ||||| ||| :||| :||| |||
Db      3347 TGGCTTCAGGTGGGACCGCCGAGGCGAGGAGCGGCTGACAGAAAGCATGCTTCGA 3406
        |||...||| ||||| ||| :||| :||| |||
Qy      999 sLeuHISpHeLYsSerHISval---TYrTYrPheAlaGlnIaGIUsErLYrThrPheG 1018
        |||...||| ||||| ||| :||| :||| |||
Db      3407 GATC---ACCGAGCCACCTCCTAGCTGTACTTACGCCCTTAGACAGGAACTACACCG 3463
        |||...||| ||||| ||| :||| :||| |||
Qy      1018 uArGTpMeGIUsnIleArGIserAla-----ThISerGIaSerArGIProH 1035
        |||...||| ||||| ||| :||| :||| |||
Db      3464 ACGCTGATGGCTGTGCTTGGCCGGCGGCGGCGGAGGAGACAGTTCTGCCGGGGCCCA 3523
        |||...||| ||||| ||| :||| :||| |||
Qy      1035 sValLeu 1037
        |||...||| ||||| ||| :||| :||| |||
Db      3524 ACTGTCTG 3531
        |||...||| ||||| ||| :||| :||| |||

RESULT 15
US-09-566-921-115
/ Sequence 115, Application US/09566921
/ Patent No. 6682888
/ GENERAL INFORMATION:
/ APPLICANT: Loring, Jeanne F.
/ APPLICANT: Tingley, Debora W.
/ APPLICANT: Edwards, Carla M.
/ FILE OF INVENTION: PA-0024 US
/ TITLE REFERENCE: GENES EXPRESSED IN ALZHEIMER'S DISEASE
/ CURRENT APPLICATION NUMBER: US/09/566,921
/ CURRENT FILING DATE: 2000-05-05
/ NUMBER OF SEQ ID NOS: 138
/ SOFTWARE: PERL Program
/ SEQ ID NO 115
/ LENGTH: 4092
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Incyte ID No. 6682888 480076.9
US-09-566-921-115

Alignment Scores:
Pred. No.: 1.88e-29 Length: 4092
Score: 424.50 Matches: 182
Percent Similarity: 40.3% Conservative: 131
Best Local Similarity: 23.4% Mismatches: 291
Query Match: 7.8% Indels: 176
DB: 3 Gaps: 28

US-09-555-342B-2 (1-1045) x US-09-566-921-115 (1-4092)
Qy      31 PROProProThrProSerGIYLYsLeuValSerIleLYsIleGlnMeLeuAspArTh 50
        |||...||| ||||| ||| :||| :||| |||
Db      196 CCGGCACCATG-CGC-----AAAACATCATGTCGTGCTGACACCATGATGAGCGAG 248
        |||...||| ||||| ||| :||| :||| |||
Qy      51 GlnGlnAlaPheGIUsnIleProGlnArGIaIaProGIYLYsValLeuLeuAspAlaValCys 70
        |||...||| ||||| ||| :||| :||| |||
Db      249 CTGGAG---TTGGCATCTCAGCCCAACACACCGGGAAGCAGTATTGACAGAGTGTG 305
        |||...||| ||||| ||| :||| :||| |||
Qy      71 AsnHISLeuAsnLeuValGlnGIYAspTYrPheGIYLeuGlnPheProAspHISLYsLYs 90
        |||...||| ||||| ||| :||| :||| |||
Db      306 AAAACATATTGGCTGAGGGAAGTTGGTCTTTGTGCTGCAAGTCAAGACACTAAAGGT 365
        |||...||| ||||| ||| :||| :||| |||
Qy      91 IleThrValTriPLeuAspLeuLeuLYsProIleValLYsGln---IleArGIrProLYs 109

```

Db 366 TTCTCCACCTGGCTGAACTCAATAAGAGTAGTCCAGAGATGGCGGAAAGAAAGC 425
Qy 110 HtHValValIValIysPheValIValIysPheProPasp--HtHrGlnLeuGln 128
Db 426 CCCCCTCTTTTAAAGTTCCGTGCCAAGTTTACCTTACCTTACGATGTGTCCGAAATGATG 485
Qy 129 GluGluLeuThrArgTrpLeuPheAlaLeuGlnValIysGlnAspLeuAlaGlnGlyArg 148
Db 486 CAGAGACATCTACCTGACCGCTGTCTTTCTGCAATGTAAGAGAGCATTTCCATATGATGAT 545
Qy 149 LeuThrCysAsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGlnSerGluIle 168
Db 546 ATTACTACCGCCGCTGAGACCGCTGTGTCTGTGCTGTGCTGTATGTGTCCAGTCTAAGTAT 605
Qy 169 GlyAspPheAspGluAlaLeuAspArgGlu--HisLeuAlaIysAsnIleTrpIlePro 187
Db 606 GCGGACTTCAATAGAAAGATGCATTAAGTCTGGCTACCTGGCCGAGAGCAAGTTGCTCCG 665
Qy 188 Gln-----GlnAspAlaLeuGlnAspIleValGln 198
Db 666 CAGAGACTCCTGGAACGCAACAACTCAACAGAGACGAGGAGGCGGATCCAGGTG 725
Qy 199 PheHisIleAsnHisIleGlyGlnThrProIleGluSerAspPheGlnLeuLeuGluIle 218
Db 726 TGGCATAGAGAACACCGTGGCATGCTCAGGAGAGATGCTGCTCGAATATCTGAAGATT 785
Qy 219 AlaArgArgLeuGlnMetTrpGlyIleArgLeuHisProAlaIysAspArgGluGlyThr 238
Db 786 GCTCAATCTGTGAGATGTGTGTGTAACCTTACGATCAAGAACAAAGAAAGGCTCA 845
Qy 239 LysIleLeuAsnLeuAlaAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIle 258
Db 846 GAGCTGTGCTGGGGGTGGATGCTGCTGGCTCTCAACATCTATAGAGCAAGATGACAGCTA 905
Qy 259 Asn-----AlaPheAsnTrpAlaLysValArgLysLeuSerPheLysArgLysArg 275
Db 906 ACCTCCCAAGATAGGCTTCCCTGGAGTGAATACAGAAATCTCTTCAATGATTAAGAAA 965
Qy 276 PheLeuIle-----LysLeuArgProAspAlaAsnSerAlaTrpGlnAspThr 291
Db 966 TTGTCTCTCAAGCCCATTTGACAAAGAACCCCGGAC----- 1001
Qy 292 LeuGlnPheLeuMetAlaSerArgAspPheCysLysSerPheTrpLysIleCysValGln 311
Db 1002 TTGCTCTTTATGCTCCCGGCTGCGGATTAACAGCGGATCTTGGCTGTGTGATGGG 1061
Qy 312 HisHisAlaPhePheArgLeuPheGlnPheGlnPheProLysProValLeuPhe 331
Db 1062 AACCATGAATATATCATG-----CGCGCTGCAAGCCTGATTAACATTAAGGTG 1109
Qy 332 SerArgLysSerSerPheArgPheSerCysArgThrGlnLysGlnValLeuAspTrpVal 351
Db 1110 CAGAGATGTAAGGACAGCCCGGAGAGAGAACCAAGAGAGAGAGAGCGTGTATG 1169
Qy 352 LysGlnGluGlyHisLysLysValGlnPheGlnArgLysHisSerLys----- 367
Db 1170 CTGAAATAAGAGAAAGAACCGTGAATGCGACAGAGAAAGAGAAAGAAAGATTGAACGG 1229
Qy 368 -----HisSerLysSerLysLeuAlaSerGlnProThrGlnLeuAsn 382
Db 1230 GAGAAAGAGAGAGCTGATGAGAGAGCTGAAGCATGACAGAAACAGTAAGAAAGCTCAG 1289
Qy 383 SerGlnValLeuGlnSerGlnGlnSerThrSerLeuThrPheGly--GluGlyAla 401
Db 1290 CAAAGATGGAAGAACAGACCGGTAGGAGCTCTGGAACCTTGAGCA--GGAACGGAAGCGTGC 1348
Qy 402 GluSerProGly-----GlyGlnSerCysArgArgGly-----LysGlnProLys 416
Db 1349 CCAAGAGGAGGCTTAAAGAGCTGGCCAAAGAGAGCTCAAGAGAGCTAAAGAGGCAAGAGCC 1408
Qy 417 ValSerIleGlyLysGluProGlySerHisProSerProAlaProArgArgProAlaGly 436
Db 436

Db 1409 CTGTGTGACAGGCTCCCGGGA-----CCAGAAAAAGACTCAAGAACAGTGGCTTTGA 1462
Qy 437 AsnLysGlnAlaAspGly-----AlaAlaSerAlaProThrGlnGlu--GluGln 452
Db 1463 AATGGCAGAGCTGACAGCTGAAATCTCCAGCTGAGAGATGGCCGACAGAAAGAGAGAG 1522
Qy 452 uGlnValVal-----LysAspArgThrGlnGlnSerLysProGlnProProGln 468
Db 1523 TAGAGCTGTGAGGTGACAGAGAGGCCAGATGTGTACAGAAAGACTTGAGAGAGACCGG 1582
Qy 468 nProSerThrGlySerLeuThrGlySerProHisLeuSerGlnLeuSerValAsnSerGln 488
Db 1583 TGTGAGCTGAAGAGCTGCCATGATGACACTATGTGGAGAGAGCTGTGAGATATAGA 1642
Qy 488 nGlyGlyValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAspThrLysGln 1643
Db 1643 G----- 1643
Qy 508 nAlaSerProLeuIleSerProLeuLeuAsnGlnAlaCysProArgThrAspAspGln 528
Db 1644 -----GATGAGCA 1651
Qy 528 uAspGlnGlyArgArgLysArgPheProThrAspLysValaTrpPheIleAlaLysGlnVal 548
Db 1652 GGATGAGAAATGGGGCAGAGCTAGTGTGACCTACGCGGTGATGCTTATGCGCAAGAGACCG 1711
Qy 548 IserThrThrGlnArgThr-----TrpLeuLysAspLeuGln 560
Db 1712 CAGTGAAGAGAAAGTACACATGAGGACAGAGAAAGATGAGCGTGTGCAAGAGACCTGAA 1771
Qy 560 uValIleThrSerTrpPheGlnSerThrValSerLysGlnAspAlaMetProGlnAlaLe 580
Db 1772 GGCCCTCACTCG-----GAGCTGGCCAAATGCCAGAGATGAGTCCAGAAAGACTGC 1822
Qy 580 uLysSerLeuIlePheProAsnPheGlnProLeuHisLysPheHisThrAsnPheLeu 600
Db 1823 CAATGACATGATCCATGCTGGAACATGAGACTGGGCGGAGACAAATTAACAGACCTTGGG 1882
Qy 600 uGlnIleGlnGlnArgLeuAlaLeuTrpGlnGlyArgSerAsnAlaGlnIleArgAspTrp 620
Db 1883 CCAAGATCCGGCAG-----GGCAACACCAAGACGCGCATTAAGAGAAAT 1924
Qy 620 rGlnArgIle-----GlyAspValMetLeuLysAsnIleGlnGln 633
Db 1925 TAGCTATAGTATGAGGACACCCAGCCTGTAGGAGACCCCTCTTCTTCCCTTGTCCCC 1984
Qy 633 yMetLysHisLeuAlaAlaHisIleLeuTrpLysHisSerGlnAlaLeuGlnAlaLeuGln 653
Db 1985 ACACTCTTACAC--CTTAACCTACCTA--ACTCATCT--GTGCTGAGGACACTTAAG 2037
Qy 653 nGlyIleLysSerSerArgArgLeuGlnAsnPheCysArgAspPheGlnLeuGlnLys-- 672
Db 2038 AGCAGCCCTGAGACTA-----TCCCAAGACTTTAAATGATGCCATGGG 2079
Qy 672 ----- 672
Db 2080 ACCAAACTAGCCCTTAGCCCCCAACCACTTCCCTGGCAAAATGATGCTCATATGG 2139
Qy 673 -----ValCysIleTrpLeuProLeuAsnThrPhe----- 681
Db 2140 TGGCAATGAAACCTCTTCTCTTCTGTGTCTCATTAAGATCTGTATGAGCTAGAAATTCCT 2199
Qy 682 -----LeuLeuArgProLeuHisArgLeuMet 690
Db 2200 ACTTCTCAAGCTTAAGATCTTTCATCTGATTTTGGCAAAATGCCCTTACACTTACTGT 2259
Qy 690 HisIleTrp--LysGlnValLeuGlnArgLysCysLysHisHisProPro 705
Db 2260 GTCTATGGGAGTCAAGTGTGAGAGTGTGGAGACTAGCTCCCTCT 2308

Search completed: May 20, 2006, 03:27:44
Job time : 548 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 20, 2006, 03:21:43 ; Search time 2764 Seconds
(without alignments)
9968.476 Million cell updates/sec

Title: US-09-555-342B-2

Perfect score: 5463
Sequence: 1 MGIEORPPRPGSRGAPENS.....SATSSARPHVLSKESLVY 1045

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 18992170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing filter 45 summaries

Command line parameters:

-MODE=frame+ .p2n.model -DEV=xlh
-Qc/abss/abssweb.spool/US0955542/rnuc 19052006 121509 13438/app_query.fasta_1
-DB=Published Applications NA Main -QPMT=fastcap -SUFFIX=p2n.rnpsm
-MINMATCH=0.1 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blomsu62 -TRANS=human0.cdt -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=pcp -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02h
-USER=US0955542 @CGN 1.1 1675 @runat 19052006 121509 13438 -NCPU=6 -ICPU=3
-NO MMP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Published Applications NA Main:
1: /EMC Celerra SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC Celerra SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC Celerra SIDS3/ptodata/2/pubpna/US09_PUBCOMB.seq:*
4: /EMC Celerra SIDS3/ptodata/2/pubpna/US09_PUBCOMB.seq:*
5: /EMC Celerra SIDS3/ptodata/2/pubpna/US09_PUBCOMB.seq:*
6: /EMC Celerra SIDS3/ptodata/2/pubpna/US10_PUBCOMB.seq:*
7: /EMC Celerra SIDS3/ptodata/2/pubpna/US10_PUBCOMB.seq:*
8: /EMC Celerra SIDS3/ptodata/2/pubpna/US10_PUBCOMB.seq:*
9: /EMC Celerra SIDS3/ptodata/2/pubpna/US10_PUBCOMB.seq:*
10: /EMC Celerra SIDS3/ptodata/2/pubpna/US10_PUBCOMB.seq:*
11: /EMC Celerra SIDS3/ptodata/2/pubpna/US10_PUBCOMB.seq:*
12: /EMC Celerra SIDS3/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /EMC Celerra SIDS3/ptodata/2/pubpna/US11_PUBCOMB.seq:*
14: /EMC Celerra SIDS3/ptodata/2/pubpna/US11_PUBCOMB.seq:*
15: /EMC Celerra SIDS3/ptodata/2/pubpna/US11_PUBCOMB.seq:*
16: /EMC Celerra SIDS3/ptodata/2/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5463	100.0	3442	7	US-10-159-563-122

2	5459	99.9	4687	9	US-10-921-707-18	Sequence 18, App1
3	5431	99.4 <td>3187</td> <td>10<th>US-10-450-763-564</th><th>Sequence 564, App</th></td>	3187	10 <th>US-10-450-763-564</th> <th>Sequence 564, App</th>	US-10-450-763-564	Sequence 564, App
4	2941.5	53.8 <td>3997</td> <td>7<th>US-10-172-118-1536</th><th>Sequence 1536, App</th></td>	3997	7 <th>US-10-172-118-1536</th> <th>Sequence 1536, App</th>	US-10-172-118-1536	Sequence 1536, App
5	2941.5	53.8 <td>3997</td> <td>8<th>US-10-342-887-1536</th><th>Sequence 1536, App</th></td>	3997	8 <th>US-10-342-887-1536</th> <th>Sequence 1536, App</th>	US-10-342-887-1536	Sequence 1536, App
6	2941.5	53.8 <td>3997</td> <td>9<th>US-10-802-432-16</th><th>Sequence 16, App1</th></td>	3997	9 <th>US-10-802-432-16</th> <th>Sequence 16, App1</th>	US-10-802-432-16	Sequence 16, App1
7	2498	45.7 <td>3094</td> <td>3<th>US-09-925-297-220</th><th>Sequence 220, App</th></td>	3094	3 <th>US-09-925-297-220</th> <th>Sequence 220, App</th>	US-09-925-297-220	Sequence 220, App
8	1006.5	18.4 <td>1718</td> <td>3<th>US-09-764-868-51</th><th>Sequence 51, App1</th></td>	1718	3 <th>US-09-764-868-51</th> <th>Sequence 51, App1</th>	US-09-764-868-51	Sequence 51, App1
9	913.5	16.7 <td>1747</td> <td>13<th>US-11-097-143-3041</th><th>Sequence 3041, App</th></td>	1747	13 <th>US-11-097-143-3041</th> <th>Sequence 3041, App</th>	US-11-097-143-3041	Sequence 3041, App
10	905.5	16.6 <td>1220</td> <td>13<th>US-11-097-143-743</th><th>Sequence 743, App</th></td>	1220	13 <th>US-11-097-143-743</th> <th>Sequence 743, App</th>	US-11-097-143-743	Sequence 743, App
11	891.5	16.3 <td>4150</td> <td>13<th>US-11-097-143-3040</th><th>Sequence 3040, App</th></td>	4150	13 <th>US-11-097-143-3040</th> <th>Sequence 3040, App</th>	US-11-097-143-3040	Sequence 3040, App
12	748.5	13.7 <td>3309</td> <td>8<th>US-10-755-889-499</th><th>Sequence 499, App</th></td>	3309	8 <th>US-10-755-889-499</th> <th>Sequence 499, App</th>	US-10-755-889-499	Sequence 499, App
13	748.5	13.7 <td>3309</td> <td>10<th>US-10-956-157-1800</th><th>Sequence 1800, App</th></td>	3309	10 <th>US-10-956-157-1800</th> <th>Sequence 1800, App</th>	US-10-956-157-1800	Sequence 1800, App
14	744	13.6 <td>4543</td> <td>8<th>US-10-152-119A-2048</th><th>Sequence 2048, App</th></td>	4543	8 <th>US-10-152-119A-2048</th> <th>Sequence 2048, App</th>	US-10-152-119A-2048	Sequence 2048, App
15	744	13.6 <td>4543</td> <td>16<th>US-11-036-196-2048</th><th>Sequence 2048, App</th></td>	4543	16 <th>US-11-036-196-2048</th> <th>Sequence 2048, App</th>	US-11-036-196-2048	Sequence 2048, App
16	743.5	13.6 <td>3387</td> <td>8<th>US-10-755-889-397</th><th>Sequence 397, App</th></td>	3387	8 <th>US-10-755-889-397</th> <th>Sequence 397, App</th>	US-10-755-889-397	Sequence 397, App
17	735.5	13.5 <td>14917</td> <td>3<th>US-09-909-667B-11</th><th>Sequence 11, App1</th></td>	14917	3 <th>US-09-909-667B-11</th> <th>Sequence 11, App1</th>	US-09-909-667B-11	Sequence 11, App1
18	724	13.3 <td>3620</td> <td>6<th>US-10-084-817-280</th><th>Sequence 280, App</th></td>	3620	6 <th>US-10-084-817-280</th> <th>Sequence 280, App</th>	US-10-084-817-280	Sequence 280, App
19	724	13.3 <td>3620</td> <td>9<th>US-10-921-707-17</th><th>Sequence 17, App1</th></td>	3620	9 <th>US-10-921-707-17</th> <th>Sequence 17, App1</th>	US-10-921-707-17	Sequence 17, App1
20	724	13.3 <td>4336</td> <td>6<th>US-10-171-581-5</th><th>Sequence 5, App1</th></td>	4336	6 <th>US-10-171-581-5</th> <th>Sequence 5, App1</th>	US-10-171-581-5	Sequence 5, App1
21	724	13.3 <td>4336</td> <td>7<th>US-10-159-563-406</th><th>Sequence 406, App</th></td>	4336	7 <th>US-10-159-563-406</th> <th>Sequence 406, App</th>	US-10-159-563-406	Sequence 406, App
22	713	13.1 <td>4922</td> <td>9<th>US-10-357-930-30232</th><th>Sequence 30232, App</th></td>	4922	9 <th>US-10-357-930-30232</th> <th>Sequence 30232, App</th>	US-10-357-930-30232	Sequence 30232, App
23	712.5	13.0 <td>6263</td> <td>10<th>US-10-756-149-4222</th><th>Sequence 4222, App</th></td>	6263	10 <th>US-10-756-149-4222</th> <th>Sequence 4222, App</th>	US-10-756-149-4222	Sequence 4222, App
24	712.5	13.0 <td>6263</td> <td>10<th>US-10-765-700-57</th><th>Sequence 57, App1</th></td>	6263	10 <th>US-10-765-700-57</th> <th>Sequence 57, App1</th>	US-10-765-700-57	Sequence 57, App1
25	703	12.9 <td>3643</td> <td>9<th>US-10-772-636-39</th><th>Sequence 39, App1</th></td>	3643	9 <th>US-10-772-636-39</th> <th>Sequence 39, App1</th>	US-10-772-636-39	Sequence 39, App1
26	703	12.9 <td>3643</td> <td>16<th>US-11-288-493-39</th><th>Sequence 2568, App</th></td>	3643	16 <th>US-11-288-493-39</th> <th>Sequence 2568, App</th>	US-11-288-493-39	Sequence 2568, App
27	700	12.8 <td>2882</td> <td>10<th>US-10-764-420-2568</th><th>Sequence 2568, App</th></td>	2882	10 <th>US-10-764-420-2568</th> <th>Sequence 2568, App</th>	US-10-764-420-2568	Sequence 2568, App
28	699.5	12.8 <td>5897</td> <td>13<th>US-11-097-143-24557</th><th>Sequence 24557, App</th></td>	5897	13 <th>US-11-097-143-24557</th> <th>Sequence 24557, App</th>	US-11-097-143-24557	Sequence 24557, App
29	695.5	12.7 <td>4125</td> <td>6<th>US-10-044-090-620</th><th>Sequence 620, App</th></td>	4125	6 <th>US-10-044-090-620</th> <th>Sequence 620, App</th>	US-10-044-090-620	Sequence 620, App
30	691.5	12.7 <td>7982</td> <td>13<th>US-11-097-143-742</th><th>Sequence 742, App</th></td>	7982	13 <th>US-11-097-143-742</th> <th>Sequence 742, App</th>	US-11-097-143-742	Sequence 742, App
31	690.5	12.6 <td>2916</td> <td>13<th>US-11-097-143-15944</th><th>Sequence 15944, App</th></td>	2916	13 <th>US-11-097-143-15944</th> <th>Sequence 15944, App</th>	US-11-097-143-15944	Sequence 15944, App
32	685.5	12.5 <td>3470</td> <td>3<th>US-11-097-143-24407</th><th>Sequence 24407, App</th></td>	3470	3 <th>US-11-097-143-24407</th> <th>Sequence 24407, App</th>	US-11-097-143-24407	Sequence 24407, App
33	675	12.4 <td>393</td> <td>3<th>US-09-918-995-8794</th><th>Sequence 8794, App</th></td>	393	3 <th>US-09-918-995-8794</th> <th>Sequence 8794, App</th>	US-09-918-995-8794	Sequence 8794, App
34	639	11.7 <td>506</td> <td>3<th>US-09-967-768A-104</th><th>Sequence 104, App</th></td>	506	3 <th>US-09-967-768A-104</th> <th>Sequence 104, App</th>	US-09-967-768A-104	Sequence 104, App
35	639	11.7 <td>506</td> <td>10<th>US-10-843-614A-6249</th><th>Sequence 6249, App</th></td>	506	10 <th>US-10-843-614A-6249</th> <th>Sequence 6249, App</th>	US-10-843-614A-6249	Sequence 6249, App
36	626	11.5 <td>3166</td> <td>9<th>US-10-370-715B-537</th><th>Sequence 537, App</th></td>	3166	9 <th>US-10-370-715B-537</th> <th>Sequence 537, App</th>	US-10-370-715B-537	Sequence 537, App
37	626	11.5 <td>3398</td> <td>8<th>US-10-302-172-201</th><th>Sequence 201, App</th></td>	3398	8 <th>US-10-302-172-201</th> <th>Sequence 201, App</th>	US-10-302-172-201	Sequence 201, App
38	609.5	11.2 <td>716</td> <td>3<th>US-09-764-868-475</th><th>Sequence 475, App</th></td>	716	3 <th>US-09-764-868-475</th> <th>Sequence 475, App</th>	US-09-764-868-475	Sequence 475, App
39	608.5	11.1 <td>2067</td> <td>8<th>US-10-250-613-23</th><th>Sequence 23, App1</th></td>	2067	8 <th>US-10-250-613-23</th> <th>Sequence 23, App1</th>	US-10-250-613-23	Sequence 23, App1
40	605	11.1 <td>3121</td> <td>7<th>US-10-108-260A-1532</th><th>Sequence 1532, App</th></td>	3121	7 <th>US-10-108-260A-1532</th> <th>Sequence 1532, App</th>	US-10-108-260A-1532	Sequence 1532, App
41	596.5	10.9 <td>717</td> <td>7<th>US-10-264-237-1390</th><th>Sequence 1390, App</th></td>	717	7 <th>US-10-264-237-1390</th> <th>Sequence 1390, App</th>	US-10-264-237-1390	Sequence 1390, App
42	590.5	10.8 <td>379</td> <td>3<th>US-09-960-352-154</th><th>Sequence 154, App</th></td>	379	3 <th>US-09-960-352-154</th> <th>Sequence 154, App</th>	US-09-960-352-154	Sequence 154, App
43	590	10.7 <td>2846</td> <td>8<th>US-10-415-187-17</th><th>Sequence 17, App1</th></td>	2846	8 <th>US-10-415-187-17</th> <th>Sequence 17, App1</th>	US-10-415-187-17	Sequence 17, App1
44	587	10.7 <td>5582</td> <td>15<th>US-11-000-688-366</th><th>Sequence 366, App</th></td>	5582	15 <th>US-11-000-688-366</th> <th>Sequence 366, App</th>	US-11-000-688-366	Sequence 366, App
45	586.5	10.7 <td>4356</td> <td>13<th>US-11-097-143-42194</th><th>Sequence 42194, App</th></td>	4356	13 <th>US-11-097-143-42194</th> <th>Sequence 42194, App</th>	US-11-097-143-42194	Sequence 42194, App

ALIGNMENTS

RESULT 1
US-10-159-563-122
Sequence 122, Application US/10159563
Publication No. US20040009154A1
GENERAL INFORMATION:
APPLICANT: Khan, Javed
APPLICANT: Ringner, Markus
APPLICANT: Peterson, Markus
APPLICANT: Melzer, Paul
TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
FILE REFERENCE: 11613.56US11
CURRENT APPLICATION NUMBER: US/10/159, 563
CURRENT FILING DATE: 2002-12-09
PRIOR APPLICATION NUMBER: US 10/133, 937
PRIOR FILING DATE: 2002-04-25
NUMBER OF SEQ ID NOS: 444
SOFTWARE: PatentIn version 3.1
SEQ ID NO 122
LENGTH: 3442
TYPE: DNA
ORGANISM: Homo sapiens
US-10-159-563-122
Alignment Scores:

Pred. No.: 0 Length: 3442
 Score: 5453.00 Matches: 1045
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 7 Gaps: 0

US-09-555-342B-2 (1-1045) x US-10-159-563-122 (1-3442)

QY 1 MetGlyGluIleGluGlnAtrProThrProGlySerArgLeuGlyAlaProGluAenSer 20
 DB 49 ATGGAGAAATAGAGAGAGCGGACCCAGATCAAGACTGGGGCCCGGAAATTCG 108
 QY 21 GYIISerThrLeuGluAtrGlyGlnLysProProProThrProSerGlyLysLeuVal 40
 DB 109 GGGATCAgTtACCTTGGAACTGGACAGAGAGCCGCCCAACCTTCAGAGAAATCTGTG 168
 QY 41 SerIleLysIleGlnMetLeuAspSerThrGlnGluAlaPheGluValProGlnArgAla 60
 DB 169 TCCATCAAAATCCAGATGCTGGATGACACCAGAGGCACTTGAAGTTCCAAAGAGCT 228
 QY 61 ProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGluGlyAspTyr 80
 DB 229 CCGGGAGAGGTGCTGTGGATGCAAGTTGCAACCACTCAACTCGTGGAGAGGTGACTAT 288
 QY 81 PheGlyLeuGluPheProAspHisLysLysIleThrValTyrLeuAspLeuLeuLysPro 100
 DB 289 TTTGGCTCGAGTTTCTGTGATCAAAAGATCAAGGTGGCTGGATCTCTTAAACCC 348
 QY 101 IleValLysGlnIleArgArgProLysIleValIleValLysPheValLysPhePhe 120
 DB 349 ATTGTGAACAGATTGAAGGCCAAAGCACGTTGTTTAAGTTTGGTGAATTTCTT 408
 QY 121 ProProAspHisThrGlnLeuGlnGluLeuThrArgTyrLeuPheAlaLeuGlnVal 140
 DB 409 CCGCTGACCAACACAACTCCAGAGAACTCAAGAGTAACTGTTGGCTGACAGTG 468
 QY 141 LysGlyAspLeuLysGlnGlyArgLeuThrCysAsnAspThrSerAlaAlaLeuLeuIle 160
 DB 469 AAGCAGAGACTTGGCTCAAGGAGGTTGACGTGTATACACAGCGGACTCTTGATTT 528
 QY 161 SerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgGluHisLeu 180
 DB 529 TCACACATTTGTGCATCTGAGATTTGGGATTTTGAATGAGCTTGGACAGAGCACTTA 588
 QY 181 AlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGluLysPheIleValGluPheHis 200
 DB 589 GCAAAATAATTAATACATACCTCAGCAAGACGCACTAGAGAGCAAAATCGTGAATTTGCAC 648
 QY 201 HisAsnHisIleGlnThrProAlaGluSerAspPheGlnLeuLeuGluIleAlaArg 220
 DB 649 CATACACACTTGGACAAACACAGCGAATTCAGATTTCCAGCTCCAGAGATTTGCCCT 708
 QY 221 ArgLeuGlnMetTyrGlyIleArgLeuHisProAlaLysAspArgGluIleThrLysIle 240
 DB 709 CGGCTGAGAGTGTATGGAAATCCGGTTTGCACCCGGCCAAAGACAGGAGGCAAGAGATC 768
 QY 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAla 260
 DB 769 AATCTGGCTCTGGCAACAGGGAATCTAGTGTTCAGGGGTTTCACTAAGATCAAGGCC 828
 QY 261 PheAsnThrAlaLysValaLysLysLeuSerPheLysArgLysArgPheLeuIleLysLeu 280
 DB 829 TTTCACTGGGCGCAAGGCGCGAGCTGACTTCAGAGAGAGAGGCTTTTCACTCAAGCTC 888
 QY 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAlaSerArgAsp 300
 DB 889 CGGCGAGATGCAATAGTGTGCTACAGAGATACCTTGAATTTCTGATGGCCAGTGGGAT 948
 QY 301 PheCysLysSerPheTyrLysIleCysValGluHisIleAlaPhePheArgLeuPheGlu 320
 DB 949 TTCTGCAAGTCTTTGGAAATCTGTGTGAACATCATGCTTTTGAACCTTTTGA 1008

QY 321 GluProLysProLysProLysProValLeuPheSerArgGlySerSerPheArgPheSer 340
 DB 1009 GAGCCCAACCAAGGCCAAGCCGCTCTTTAGCCGGGGGTCATCTTCCGTTCACT 1068
 QY 341 GlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyValLysLysValGln 360
 DB 1069 GGTGGAGTCAAGAGCAGGTTCTCGACTATGTAAAGAGAGGACATAGAAGGTGACG 1128
 QY 361 PheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
 DB 1129 TTTGAAGAGAGACAGAAAGATTCACTTATCCGAGGCTTGCTTCAAGCTTACAGAA 1188
 QY 381 LeuAsnSerGluValLeuGlnGlnSerGlnGlnSerThrSerLeuThrPheGlyGluGly 400
 DB 1189 CTGAATTCGGAAGTGTGGAGAGCTTCAAGAGAGACCAAGCTTCACTTGGAGAAAGT 1248
 QY 401 AlaGluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGly 420
 DB 1249 GCCGAATCTCAAGGGGGCAGAGCTGCGGCGAGAAAGAACCGAAAGTTTCCGGCGGG 1308
 QY 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAla 440
 DB 1309 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAGAGCCCGGGGTAAACAGAGCGG 1368
 QY 441 AspGlyAlaAlaSerAlaProThrArgGluGluGluGluValIleLysAspArgThrGln 460
 DB 1369 GACGAGCGCGCTCGGCGGCCACGAGAGAGAGAGAGAGAGTGTTAAGATAGACCCAG 1428
 QY 461 GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeu 480
 DB 1429 CAGAGTAAACCTCAGCCCCCGCAGCCACAGCAGGCTCTCTGACTGCGAGTCTCACTT 1488
 QY 481 SerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro 500
 DB 1489 TCCAGCTGTCTGTGAATCTGAGGGGAGAGTGGCCCTTGCAAGCTGATCTGTCTCC 1548
 QY 501 AsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuLeuAsnArgGln 520
 DB 1549 AACCTGAGCCCGGACACCAAGAGAGGCTCTCTCTTATCAGCCGCTGTGATAGACAG 1608
 QY 521 AlaCysProArgThrAspAspGluAspGluGlyArgArgLysArgPheProThrAspLys 540
 DB 1609 GCCTGCCCCCGGACGACATGAGATGAGGGCCGAGAGAAAGATTTCCACGTGTA 1668
 QY 541 TyrThrPheIleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAspLeuGlu 560
 DB 1669 GCCTACTTCAATGCTAAGAGAGTGTTCACCAAGAGCAACATATGTGAAGATCTCGAA 1728
 QY 561 ValIleThrSerTyrPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeu 580
 DB 1729 GTTATCACTTCGTCGTTTCAAGACACATGAGCAAAAGAGAGAGCCCATGCGGAAAGCTG 1788
 QY 581 LysSerLeuIlePheProAsnPheGluProLeuHisLysArgPheHisThrAsnPheLeuLys 600
 DB 1789 AAAAGTCAATATTCGGAATTTGAACCTTGGCAAAATTTCAATCAATTTTCCAAAG 1848
 QY 601 GluIleGluGlnArgLeuAlaLeuTyrGluGlyArgSerAsnAlaGlnIleArgAspTyr 620
 DB 1849 GAATATTGAGCAAGACTTGCTGCTGGAGAGGCGCTCAATATCCAAATCAAGATTTAC 1908
 QY 621 GlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaAlaHis 640
 DB 1909 CAAGAGATCGGAGATGTCATGTGAAGAACATTCAGGGCATGAAGCACTGGCGGCTCAC 1968
 QY 641 LeuTyrLysHisSerGluAlaLeuGluAlaLeuGluLysArgGlyLysSerSerArgArg 660
 DB 1969 CTGTGGAACACAGCAGAGGCTTGGAGGCTCTGGAAGATGAAGATTCAGAGCTCCGAGCGG 2028
 QY 661 LeuGluAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuProLeuAsnThr 680
 DB 2029 CTGGAGAACTTCTGCGAGAGCTTTGAGCTGCGAGAGGTGTGTTAACTTCCGCTCAACACC 2088
 QY 681 PheLeuLeuArgProLeuHisArgLeuMetHisTyrLysGlnValLeuGluArgLeuCys 700

```

Db      2089 TTCTCCGCGGCGCAGCTGCAACGGCTCATGCACTAAGAGAGGCTCTGAGACGGCTGTC
Qy      701  LysHSHSPROFROSERHIALAAPHENARGPYBAAGLALALAEULALGLUTLE
Db      2149 AAACACCAACCCGCGAGCCAGCCGACTTCAGGGACTGCGCAGCGCTTTGGCAGAGATTC
Qy      721  ThrGluMetValAlaGluLeuHISGLYThrMetIleLysMetGluAsnPheGluLysLeu
Db      2209 ACCGAGATGGTGGACAGCTCAGCGTACGATGATCAAGATGAGAAATTTCCAGAGCTG
Qy      741  HisGluLeuLysLysAspLeuIleGlyIleAspAsnLeuValAlaProGlyArgGluPhe
Db      2269 CACGAACCTCAAGAAAGATTGATTGAGCATTCATGATGTTGGTTCCGGGAAGGAGTTC
Qy      761  IleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArgMetPhePhe
Db      2329 ATCCGTCTGGGACCTCAGCAAGCTCTCGGGGGAAGGGGCTCCAGCAGCGCATGTTCTTC
Qy      781  LeuPheAsnAspValIleLeuLysLysThrSerArgGlyLeuThrAlaSerAsnGlnPheLys
Db      2389 CTGTTCAACGACGCTCTGCTATACAGACCGGGGGCTGACGGCTCCCAATCAGTTTAAA
Qy      801  ValHisGlyGlnLeuProLeuLysGlyMetThrIleGluGlnSerGluAspGluTrpGly
Db      2449 GTCCACGGGAGCTCCGCTCTAGGATGAGCATGAGTGAAGAGAGAGAGAGAGAGTGGGGG
Qy      821  ValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaLysSer
Db      2509 GTGCGCCACCTGCGACCTCCGCGGGCCAGCGGCGCATCATCATCTGGCCGACGTTCT
Qy      841  ArgSerGluMetGluLysTrpValGluAspIleGlnMetAlaIleAspLeuAlaGluLys
Db      2569 CGGTCCAGATGAGAGTGGGTTGAGGACATCCAGATGGCCATTGACCTGGCCGAGAGG
Qy      861  SerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLysSerProAsp
Db      2629 AGCAGCAGCCCGCCCGCTGAGTTCTGCGCAGCAGCCCGCTGCAACAATCCCTGAT
Qy      881  GluAlaThrAlaAlaAspGlnGlnSerGluAspLeuSerAlaSerArgThrSerLeu
Db      2689 GAGGCCACCGCGGCTGACAGAGATCAGAGATACCTGAGCGCTCGCGCACATCTGCTG
Qy      901  GluArgGlnAlaProHisArgGlyAsnThrMetValHisValCysTrpHisArgAsnThr
Db      2749 GAGCGCCAGGCGCCGCGCAGCGGCAACAATGGTGCAGTGTCTGGCACCGCAACCC
Qy      921  SerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerGlyAsnLeuLeu
Db      2809 AGCGTCCATGGTGAAGCTTCAAGATCGCAGTGAAGATCAGTTGTCTGGAATCTGCTG
Qy      941  ArgLysPheLysAsnSerAsnGlyTrpGlnLysLeuTrpValValPheThrAsnPheCys
Db      2869 AGGAAATTCAAAACACGAAACGGGTGGCAGAAAGCTGTGGGTGTTCAAAATCTTCTGC
Qy      961  LeuPhePheLysSerHisGlnAspAsnHisProLeuAlaSerLeuProLeuLeuGly
Db      2929 CTGTTCTTCAACAATCAACACAGAGAACATATCCCTTGCACACCTGCTCTCTGCGC
Qy      981  TyrSerLeuThrIleProSerGluSerGluAsnIleGlnLysAspTrpValPheLysLeu
Db      2989 TACTCGCTCAACATCCCTCTGAGTCCGAGAAACATCCAGAAAGCTACGTCTCAACTG
Qy      1001 HisPheLysSerHisValLysLysTrpPheArgLysGluSerGlyLysPheGluArgTrp
Db      3049 CACTTCAAGTCCCACTGCTACTTCAAGGCGGAAAGCGAGTCAAGTTTCAAAAGGTGG
Qy      1021 MetGluValIleAspSerAlaThrSerSerAlaSerArgProHisValIleLeuSerHis
Db      3109 ATGGAAGTATCCGACGTCACAGCTCTGCTTGGGAGCCCAAGTGTGAGCCACAAA
Qy      1041 GluSerLeuValLys 1045

```

```

Db      3169 GAGTCTCTGTGTAT 3183
RESULT 2
US-10-921-707-18
; Sequence 18, Application US/10921707
; Publication No. US2005003447A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILDMAN, Jennifer L.
; APPLICANT: BANDMAN, Olga
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0594 PCT
; CURRENT APPLICATION NUMBER: US/10/921,707
; CURRENT FILING DATE: 2004-08-19
; PRIOR APPLICATION NUMBER: US/09/786,797
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: 09/156,470; unassigned: 60/131,321
; PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 18
; LENGTH: 4687
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1320252
US-10-921-707-18
Alignment Scores:
Pred. No.: 0 Length: 4687
Score: 5459.00 Matches: 1044
Percent Similarity: 99.9% Conservative: 0
Best Local Similarity: 99.9% Mismatches: 1
Query Match: 99.9% Indels: 0
Gaps: 0
US-09-555-342B-2 (1-1045) x US-10-921-707-18 (1-4687)
Qy      1  MetGlyLysIleGluGlnArgProThrProGlySerArgLeuGlyAlaProGluAsnSer
Db      90  ATGGAGAAATTAAGAGAGAGCGCCAGCCAGATCAAGACTGGGGGCGCCGGAAAAATTGG
Qy      21  GlyIleSerThrLeuGluArgGlyGlnLysProProProThrProSerGlyLysLeuVal
Db      150  GGGATCAGTACCTTGGAACTGGACAGAAACCCGCCCAACACTTCAGGAAAACTGCTG
Qy      41  SerIleLysIleGlnMetLysAspAspThrGlnGluAlaPheGluValAlaProGlnArgAla
Db      210  TCCATCAAAATCCAGATGCTGAGTGAACACCCAGAGAGCAATTTGAAGTTCCACAAAGAGCT
Qy      61  ProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGluGlyAspTrp
Db      270  CTGGGAGAGGTGCTGCTGAGTGAATGCAACAGTTTGCAACCACTCAACTCGTGAAGGACTAT
Qy      81  PheGlyLeuGluPheProAspHisLysLysLysIleThrValTrpLeuAspLeuLysPro
Db      330  TTGGGCTCCAGATTCTCTGATCACAAAAGATACGGTGTGGTGTCTCTTAAGAACTT
Qy      101  IleValIleGlnIleArgAspProLysHisValValLysPheValValLysPhePhe
Db      390  ATTGTGAACAGATTGAAGGCCAAAGACAGCTGTGTGTTAAGTTGTGTGTAATTTCTT
Qy      121  ProProAspHisThrGlnLeuGlnGluLeuThrArgTrpLeuPheAlaLeuGlnVal

```

Dh 450 CCGCTGACCACACAACTCCAAGAACTCAAGGTAACCTGTTCCGGCTGACGGTG 509
Qy 141 LysGlaAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerLysAlaLeuLeuIle 160
Dh 510 AACGAGACTGGCTCAAGGACAGGTGACGTGTAATGACACACGACGACTCTTGATG 569
Qy 161 SerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgGluHisIleu 180
Dh 570 TCACACTTGTGCATCTGAGATTGGGGATTGTTGATGTAACCTTTGACAGAGACTTA 629
Qy 181 AlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIleValGluPheHis 200
Dh 630 GCAAAAAATTAATATCACTCTGACGAAGACGCACTGAGAGCAAAATTCGGGAATTTCC 689
Qy 201 HisAsnHisIleGlyGlnThrProAlaGlnSerAspPheGlnLeuLeuGluIleAlaArg 220
Dh 690 CATTAACCACTTGACAAACACAGCAGATCAAGATTTCAGCTCTTAAGATTTGCCCT 749
Qy 221 ArgLeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysIle 240
Dh 750 CCGCTAGAGATGTAATCGGGTTCACCCGGCCAAAGGACAGGGAGGCAAGAGATC 809
Qy 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAla 260
Dh 810 AATCTGGCCGTTCACCAACACGGGAATTTCTAGTGTTCAGGGGTTTCACTAAGATCAATGCC 869
Qy 261 PheAsnThrAlaLysValArgLysLeuSerPheLysArgLysArgPheLeuIleLysLeu 280
Dh 870 TTCACTGGGGCAAGGGCGGAAGCTGAGCTTCAAGAGGAAGGCTTTTCAACAAGCTC 929
Qy 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAlaSerArgAsp 300
Dh 930 CGGCGAGATGCCAATAGTGTGTCACAGGATACCTTGGAATTCCTGATGGCAGATCGGGAT 989
Qy 301 PheCysLysSerPheThrLysIleCysValGluHisHisAlaPhePheArgLeuPheGlu 320
Dh 990 TTCTGCAAGCTCTTCGGAATACTGTGTGAACATCATGCTCTTCAACTTTTGTGA 1049
Qy 321 GluProLysProLysProLysProValLeuPheSerArgLysSerPheArgPheSer 340
Dh 1050 GAACCCCAACCAAGCCCAAGCCGCTCTTTAAGCGGGGGATCATATTCCGATCAAGT 1109
Qy 341 GlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyHisLysLysValGln 360
Dh 1110 GGTCCGACTCAGAGGAGGTTCTCGACTATGTTAAGAAGAGAGACATAAGAGGTGCAG 1169
Qy 361 PheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
Dh 1170 TTGGAAGAAGACACAGCAAGATTCTATCTCGGAGCCTTCTCACAGCTTACAGAA 1229
Qy 381 LeuAsnSerGluValLeuGlnGlnSerGlnGlnSerThrSerLeuThrPheGluGluGly 400
Dh 1230 CTGAATTCGAAATGCTGGAGCATCTTCACAGACACAGCCTTACATTTGGAGAAGGT 1289
Qy 401 AlaGlnSerProGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGly 420
Dh 1290 GCGGAATCTCAGAGGGGCGCAGACTCGCGCGAGAAAGAACCGAAGTTTCGCGCGG 1349
Qy 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAla 440
Dh 1350 GAACCGGGGTCCAGCCGACCGGACCTGCGCGCAGAGAAAGCCCGGGGTAAACAAGCGGG 1409
Qy 441 AspGluAlaIleAspAlaProThrGlnGlnGluGlnGluValValLysAspArgGln 460
Dh 1410 GAAGAGACCCCTCGGGGCCACGAGAGAAAGAGAGGTCTGTTAAGATAGAACCCAG 1469
Qy 461 GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrArgLysProHisIleu 480
Dh 1470 CAAAGTAACCTCAGGCCCGGACGACCAAGCAAGGCTCTGACTGGCACTTCCACCTT 1529
Qy 481 SerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro 500
Dh 1530 TCCGAGCTGTCTGTGAACCTCGAGGGGGAGTGGCCCTCGCAACGTGAACCTTGTCTCCC 1589

Qy 501 AsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuLeuAsnArgGln 520
Dh 1590 AACCTGAGCCCGGACACCAAGAGAGGCTCTCCCTTGATGACGCCGCTGATGATGACAG 1649
Qy 521 AlaCysProArgThrAspAspGluAspGluGlyArgArgLysArgPheProThrAspLys 540
Dh 1650 GCCTGCCCGGACGGACGATGAGATGAGGGCCGGAGGAAGATTTCCCACTGATAA 1709
Qy 541 AlaTyrPheIleAlaLysGlnValSerThrThrGlnArgThrThrLeuLysAspLeuGlu 560
Dh 1710 GGTACTTATAGCTTAAGGAAGTCTTACACGAGCCGAACATATCTGAAGATCTCGAA 1769
Qy 561 ValIleThrSerThrPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeu 580
Dh 1770 GTTATACCTTCGTGGTTTCAAGACACAGTGAAGAAAGAGAGCCCATGCGGAAGCACTG 1829
Qy 581 LysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAsnPheLeuLys 600
Dh 1830 AAAAGTCTATATTCGCAATTTTGAACCTTTGCACAAAATTTCACTAATTTTCTCAAG 1889
Qy 601 GlnIleGlnGlnArgLeuAlaLeuTrpGluLysArgSerAsnAlaGlnIleArgAspTyr 620
Dh 1890 GAATTCGACCAACGACTTCCCTGTGGAAAGCCGCTCAATGCCCCAATCAGAGATTAAC 1949
Qy 621 GlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaHis 640
Dh 1950 CAAAGATCGCGGATGTCACTGCTGAAGACATTCAGGGCATGAAGCACTGGCGGCTCAC 2009
Qy 641 LeuTrpLysHisSerGluAlaLeuGluValAlaGluAsnGlyIleLysSerSerArgArg 660
Dh 2010 CTGTGGAAACACACGACGCTTGGAGGCCCTCGAGAAATGGAATCAAGAGCTCCCGCGG 2069
Qy 661 LeuGluAsnPheCysArgAspPheGluLeuGluLysValCysTyrLeuProLeuAsnThr 680
Dh 2070 CTGGAAGACTTCGACAGACTTGAAGCTGACAGAGGTGTGTTACTTACCGCTCAACACC 2129
Qy 681 PheLeuLeuArgProLeuHisArgLeuMetHisTyrLysGlnValLeuGluArgLeuCys 700
Dh 2130 TTCTCTCTCGGCGCACTGACCGGCTCATGACACTCAACACAGTCTCGAAGCGGCTGTG 2189
Qy 701 LysHisHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIle 720
Dh 2190 AAAACACACCCCGGACGACCCGCACTTCAGGGACTCGCGGACCCCTTTGGCAGAGATC 2249
Qy 721 ThrGluMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPheGluLysLeu 740
Dh 2250 ACGAGATGATGACAGACTCCACGCTACGATGATCAAGATGGAATTTCCAGAACTG 2309
Qy 741 HisGluLeuLysLysAspLeuIleGlyIleAspAsnLeuValValProGlyArgGluPhe 760
Dh 2310 CAGGAATCAAGAAAGATTGATTGGCAATTGAACAATCTGTGGTTCCGGGAAGGAGTTTC 2369
Qy 761 IleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArgPhePhe 780
Dh 2370 ATCCGCTCTGGGAGCTTCAGCAAGCTCTCGGGGAAGGGGCTCCAGCAGCGCAATGTTCTTC 2429
Qy 781 LeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsnGlnPheLys 800
Dh 2430 CTGTTCAAGGAGGCTCTGCTAATACAGAGCCGGGGGCTACGGCTTCAATCAAGTTAA 2489
Qy 801 ValHisGlyGlnLeuProLeuTyrGlyMetThrIleGluGlnSerGluAspGluTrpGly 820
Dh 2490 GTCCAGGGGACACTCCGCTCATGAGCAGAGATTGAAGAGGCGAAGCGAGTGGGG 2549
Qy 821 ValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAlaSerSer 840
Dh 2550 GTGCCCCACTGCTGACTCTCCGGGGCGCAGCGAGTCCATCATCTGCGCGGCAATTCCT 2609
Qy 841 ArgSerGluMetGluLysTrpValGluAspIleGlnMetAlaIleAspLeuAlaGluLys 860
Dh 2610 CGGTCCGAGATGAGAAAGTGGTTGAGACATCCAGATGACCTTACCTTGGCGGAAAG 2669

Qy 861 SerSerSerProAlaProGluPheLeuAlaSerSerProProAlaAsnLysSerProAsp 880
Db 2670 AGCAGCAGCCCCCGCCCTGAGTCTTGCCAGCAGCCCCCTTGACAAACAAGTCCCTAT 2729
Qy 881 GluAlaThrAlaAlaAspGlnGluSerGluAspAspLeuSerAlaSerArgThrSerLeu 900
Db 2730 GAAGCAGCCGGGCTGACCAAGAGTCAAGAGTACCTGAGGCTCCGCAACATCGCTG 2789
Qy 901 GluArgGlnAlaProAlaArgGluAsnThrMetValHisValCysThrPheArgAsnThr 920
Db 2790 GAGGCGCAGGCCCGCCAGCCGCGCAACAATGTGCGCTGTGCGACCGCAACCC 2849
Qy 921 SerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerGluAsnLeu 940
Db 2850 AGCGTCTCCATGCGTGAATTCAGCATCCGAGTGGAAATCAGTTGTCTGGAAACCTGCTG 2909
Qy 941 ArgLysPheLysAsnSerAsnGlyTrpGlnLysLeuTrpValPheThrAsnPheCys 960
Db 2910 AGGAAATTCAAAAACAGCAACGGGTGGCAAGAGTGTGGTGTTCACAAACTTCTGC 2969
Qy 961 LeuPhePheTrpLysSerHisGlnAspAsnHisProLeuAlaSerLeuProLeuLeu 980
Db 2970 CTGTCTCTTCAAAATCACACAGGACATCATCCCTTGCCAGCTGCTGCTGCTGCTG 3029
Qy 981 TyrSerLeuThrIleProSerGluSerGluSerGluSerGluSerGluSerGluSerGlu 1000
Db 3030 TACTCGCTCACCACTCCCTCTGAGTCCAGAAACATCCGAAAGACTACGTTCAGCTG 3089
Qy 1001 HisPheLysSerHisValTrpPheArgAlaGluSerGluTrpPheGluArgTrp 1020
Db 3090 CACTTCAAGTCCCAAGCTTACTACTTCAAGGGCGGAAACGAGTACAGTTGAAAGGTGG 3149
Qy 1021 MetGluValIleArgSerAlaThrSerSerAlaSerAlaGluProHisValLeuSerHis 1040
Db 3150 ATGGAAGGATCCCAAGTGCACCAAGCTCTGCTGCGACCCAGTGTGATGACAA 3209
Qy 1041 GluSerLeuValTyr 1045
Db 3210 GAGTCTCTTGTGTAT 3224

RESULT 3
US-10-450-763-564
; Sequence 564, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 564
; LENGTH: 3187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (49)..(1371)
; OTHER INFORMATION: 92% homologous to Homo sapiens CDEP, accession number
; OTHER INFORMATION: AB008430, Smith-Waterman Score=2090.
US-10-450-763-564

Alignment Scores:
Pred. No.: 0 Length: 3187
Score: 5431.00 Matches: 1042
Percent Similarity: 99.6% Conservative: 0

Best Local Similarity: 99.6% Mismatches: 3
Query Match: 99.4% Indels: 1
DB: 10 Gaps: 0

US-09-555-342b-2 (1-1045) x US-10-450-763-564 (1-3187)

Qy 1 MetGluValIleGlnGlnArgProThrProGluSerArgLeuGluValaProGluAsnSer 20
Db 49 ATGGAGAAATAGACAGAGGCCAAGCCAGAGATCAAGATCGGGGGCCCCGAAATATCG 108
Qy 21 GlyIleSerThrLeuGluArgGluGlnLysProProProThrProSerGluLysLeuVal 40
Db 109 GGATACAGTACTTGGAAAGTGCAGAGAGCCGCCCAACACTTCAGAGAAACTCGTG 168
Qy 41 SerIleValIleGlnMetLeuAspAspThrGlnGlnAlaPheGlnValaProGluArgAla 60
Db 169 TCCATCAAAATCCAGATGCTGTGATGACCCAGAGAGGATTTGAAATTCACAAAGAGCT 228
Qy 61 ProGluLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGluGluAspTyr 80
Db 229 CTTGGAGAGTGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 288
Qy 81 PheGluLysGluProAspHisLysLysIleThrValTrpLeuAspLeuLysPro 100
Db 289 TTTGGCTCGAGTTTCTGATCAAAAAGATCAAGGTGTGTGATCTCTTAAATCC 348
Qy 101 IleValLysGlnIleArgArgProLysHisValValValLysPheValValLysPhe 120
Db 349 ATTGTGAACAGATTGAAGGCCAAAGCAAGCTTTGTGTAAGTTGTGTGTAATTTCTT 408
Qy 121 ProProAspHisThrGlnLeuGlnGluGluLeuThrArgTrpLeuPheAlaLeuGlnVal 140
Db 409 CCGCTTACCAACAACAATCTCAAGAAAGACTCAAGAGTCTCTTGGCTGAGGTG 468
Qy 141 LysGlnAspLeuAlaGlnArgLysThrCysAsnAspThrSerAlaAlaLeuLeuIle 160
Db 469 AAGCAGAGCTTGCTCAAGGAGGTGACGTGATATGACACAGCGAGCTCTTGAAT 528
Qy 161 SerHisIleValGlnSerGluIleGluAspPheArgLysAlaLeuAspArgLysIleLeu 180
Db 529 TCACACATTGTGCAATCGAATTTGGATTTTATGTAAGCTTGGACAGAGACACTTA 588
Qy 181 AlaLysAsnLysTrpIleProGlnGlnAspAlaLeuGluAspLysIleValGluPheHis 200
Db 589 GCATAAATATATATCATCTTCAACAAGCCACTAGAGCAAAATCTGTAATTTTAC 648
Qy 201 HisAsnHisIleGlnIleThrProAlaGluSerAspPheGlnLeuLeuGluIleAlaArg 220
Db 649 CATAAACCATTTGAAACAAACAGCAGGAATCATATTTCCAGTCTTGAAGATTCGCGT 708
Qy 221 ArgLeuGluMetTrpGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysIle 240
Db 709 CGGTAGATGATGTGGAATCCGGTGTGACCCGCGCAAGGACAGAGGACAGCAAGTTC 768
Qy 241 AsnLeuAlaValaAlaAsnThrGlyIleLeuValPheGlnGluPheThrLysIleAsnAla 260
Db 769 AATCTGGCGTTGCCAACAGGGAATTTAGTGTTCAGGCTTTCACATAAGATAATGCC 828
Qy 261 PheAsnTrpAlaLysValArgLysLeuSerPheLysArgLysArgPheLeuIleLysLeu 280
Db 829 TTCAACTGGGCGCAAGGCGGAACTGAGCTTCAAGAGGAAGCGCTTCTCATCAAGTCC 888
Qy 281 ArgProAspAlaAsnSerAlaTrpGlnAspThrLeuGluPheLeuMetAlaSerArgAsp 300
Db 889 CGGCAATGCAATATGCTGATCAGAGATACCTTGAATTTCCGATGGCCAGTGGAGT 948
Qy 301 PheCysLysSerPheTrpLysIleCysValGluHisHisAlaPhePheArgLeuPheGlu 320
Db 949 TTTGCAAGTCTCTGGAATCTGTGTGAAATCATAGCTTCTTTAGACTTTTGAA 1008
Qy 321 GluProLysProLysProLysProValLeuPheSerArgGluSerSerPheArgPheSer 340
Db 1009 GAGCCCAAAACCAAGCCCAAGCCGCTCTTGTAGCCGGGGTGCATATTTGGTTCAGT 1068

QY 341 G|YArGThrGlnYsgInValIleuAspTyxValIlysgInGlyH|sLYsLYsValGln 360
|
Db 1069 GGTGGAGCTGAAAGCAGGTTCTGACTATCTTAAAGAGAGACATAAGAGGTGACG 1128
QY 361 PheGlnuArgLYsH|sSerLYsL|eH|sSerL|eArGserLeuAlAserGlnProThrGlu 380
|
Db 1129 TTTGAAAGAAAGCACAACAAGATTCAATTCTATCCGGAGCCTTGCTTACAGCCTTAAGAA 1188
QY 381 LeuAsnSerGluValIleuGlnGlnSerGlnGlnSerThrSerLeuThrPheGlu-Glu 400
|
Db 1189 CTGAATTCTGGAAGTGTGAGACAGTCTTCAAGAGCAACAGCCTTACATTTGAGAGAGG 1248
QY 400 YAlAsGlnSerProGlnYgInGlnSerCyArGArgGlyLYsGlnProLYsValSerAlaG1 420
|
Db 1249 TGGCGAATTTCCAGGGGGCCAGAGCTCCGGCGAGAAAGAAACCAAGAGTTCCGGCGG 1308
QY 420 YGlnProGlySerH|sProSerProAlaProArgArgSerProAlaGlyAsnLYsGlnAl 440
|
Db 1309 GAGAGCCGGGGTCCGACCCGAGCCTCGCGCGAGAGAGCCCGCGGGGTAAACAAGCAGGC 1368
QY 440 aaPpGlyAlaAlAserAlaProThrGlnGlnGlnGlnGlnValValLYsAspArgThrG1 460
|
Db 1369 GGAAGGAGCGGCTCGAGGCCCAAGAGAGAGAGAGAGAGGTGCTTAAAGATAGGACCA 1428
QY 460 nGlnSerLYsProGlnProProGlnProSerThrGlySerLeuThrGlySerProH|sLe 480
|
Db 1429 GCAAGATTAACCTCAGCCCTCGAGCCCAAGCAACAGGCTCTCTGATGCTGGAGTCTCACT 1488
QY 480 uSerGlnLeuSerValAsnSerGlnGlyLYsAlaAlaProAlaAsnValThrLeuSerPr 500
|
Db 1489 TTTCCGAGCTGCTGTGAACCTCGAGGGGGAGTGCCCTTGCCCACTGACCTTGCTGCC 1548
QY 500 aaSnLeuSerProAspThrLYsGlnAlAserProLeuL|sSerProLeuAsnAspG1 520
|
Db 1549 CAACCTAGCCCTCAGACCAAGAGAGGCTCTCTCTTATCAGCCGCTGCTGAATGACCA 1608
QY 520 nAlaCyProArgThrAspAspGlnuAspGlnGlyLYsArgLYsAspPheProThrAspLY 540
|
Db 1609 GGCTGCCCCCGAGCGAGCATAGAGATGAGGGCCGAGAGAGAGATTTCCCACTGATTA 1668
QY 540 sAlaTYrPheL|eAlaLYsGlnValSerThrThrGlnuArgThrTYrLeuLYsAspLeuG1 560
|
Db 1669 AGCGTACTTCAAGTGAAGAGATGCTACCAACGAGCAACATATCTGAAGGATCTCGA 1728
QY 560 uValL|eThSerThrPheGlnSerThrValSerLYsGluAspAlaMetProGluAlaLe 580
|
Db 1729 AGTTATCACTTGTGGTTTCAAGACACAGTGAAGAAAGAGACGCCATGCCGGAAGCACT 1788
QY 580 uLYsSerLeuL|sPheProAsnPheGlnProLeuH|sLYsPheH|sThrAsnPheLeuLY 600
|
Db 1789 GAAAGGCTCATATTTCCGAAATTTTGAACCTTTGCAACAATTTCTAAATTTTCTCAA 1848
QY 600 sGlnL|eGlnGlnArgLeuAlaLeuTYrGlnGlyLYsXSerAsnAlaGlnL|eArgAspTY 620
|
Db 1849 GGAATTTGACCAACGACTTCCCTGTGGAAAGGCCCTCAATATGCCAAATTCAGAGATTA 1908
QY 620 rGlnuArgL|eGlyAspValMetLeuLYsAsnL|eGlnGlyMetLYsH|sLeuAlaAlaH| 640
|
Db 1909 CCAAGAGATGGAGATGTCATGCTGAAGAACATTCAAGAGCATGAAGCACTGGGGGTCTCA 1968
QY 640 sLeuTYrLYsH|sSerGlnAlaLeuGlnAlaLeuGlnLYsValLYsSerSerArgAr 660
|
Db 1969 CTTGTGGAACACAGCGAGGCTTTGAGGCGCTTGGAAATGGAATCAAGAGCTCCGGCG 2028
QY 660 gLeuGlnAsnPheCyArGArgPheGlnLeuGlnLYsValCyArTYrLeuProLeuAsnTh 680
|
Db 2029 GCTGGAGAACTTCTGAGAGACTTTGAGCTGCAAGAGGTGTGTTACTTACCGCTCAACAC 2088
QY 680 rPheLeuLeuArgProLeuH|sArgLeuMetH|sTYrLYsGlnValLeuGlnuArgLeuCY 700
|
Db 2089 CTTCTCTCTGGGCACTGACCGGCTCATGCACTTAACAAGAGGTCTCTGAGCGGCTGTG 2148

QY 700 sLYsH|sH|sProProSerH|sAlaAspPheArgAspCyArGAlaAlaLeuAlaGlnL 720
|
Db 2149 CAACAACACCCGCCAGCCAGCCGACCTTCAAGGAGCTGCCAGCGGCTTTGGCAAGAT 2208
QY 720 eThrGlnMetValAlaGlnLeuH|sGlyThrMetL|eLYsMetGlnAsnPheGlnLYsLe 740
|
Db 2209 CACGAGATGTGTGACAGCTTCCACGGTACGAGATCAAGATGAAGATTTTCCAGAGACT 2268
QY 740 uH|sGlnLeuLYsAspLeuL|eGlyL|eAspAsnLeuValValProGlyArgGluPh 760
|
Db 2269 GCAAGAACTCAAAAGATTTGATTTGACATTGACAAATCTTGTTGCTCCGGAAAGGAGTT 2328
QY 760 eL|eArgLeuGlySerLeuSerLYsLeuSerGlyLYsGlyLeuGlnGlnArgMetPhePh 780
|
Db 2329 CATCCGCTGTGGCAGCTCAGCAAGCTTCGGGGGAAGGGCTCCACAGCGCATGTTCTT 2388
QY 780 eLeuPheAsnAspValLeuLeuTYrThrSerArgGlyLeuThrAlaSerAsnGlnPheLY 800
|
Db 2389 CCTGTTCAACGAGTCTCTGTATACAGAGCCGGGGGCTGACGGCTTCCAAATCAGTTTAA 2448
QY 800 sValH|sGlyGlnLeuProLeuTYrGlyMetThrL|eGlnGlnSerGlnuAspGluTYrPgl 820
|
Db 2449 AGTCCAGGCACTTCCGCTCTATGACATGACGATTTGAGAGAGAGAGACAGATGGGG 2508
QY 820 YValProH|sCyLYsLeuThrLeuArgGlyGlnArgGlnSerL|eL|eValAlaAlaSerSe 840
|
Db 2509 GGTGCCCACTGCTGACCTCCGGGGCCAGCGGCGCATCATGCTGGCGCCCAAGTTTC 2568
QY 840 rArgSerGlnMetGlnLYsTYrValGlnuAspL|eGlnMetAlaL|eAspLeuAlaGluLY 860
|
Db 2569 TCGGTCCGAGATGGAGATGGGTTGAGACATCCAGATGGCCATTGACCTTGGCGAGAA 2628
QY 860 sSerSerSerProAlaProGlnPheLeuAlaSerSerProProAspAsnLYsSerProAs 880
|
Db 2629 GAGCAGAGCCCGCCCTGAGTTCCTGGCCAGCAGCCCTTGAACAACAGTCCCTTGA 2688
QY 880 pGlnAlaThrAlaAlaAspGlnGlnuSerGlnuAspAspLeuSerAlaSerArgThrSerLe 900
|
Db 2689 TGAAGCCACCGGGGCTGACAGAGATCAAGAGATGACCTGAGCGCTCCGCGCAATCGCT 2748
QY 900 uGlnuArgGlnAlaProH|sArgLYsAsnThrMetValH|sValCyArTYrH|sArgAsnTh 920
|
Db 2749 GAGGCCAGGCCCCCAGCCGAGCAACCAATGATGCAAGTGTGTGGACCGCAACAC 2808
QY 920 rSerValSerMetValAspPheSerL|eAlaValGlnuAsnGlnLeuSerGlyAsnLeuLe 940
|
Db 2809 CAGCGTCTCATGTGTGACTTCAAGCATGCCAGTGGGAATCAGTTGTCTGGAACCTTGT 2868
QY 940 uArgLYsPheLYsAsnSerAsnGlyTYrPgluLYsLeuTYrValValPheThrAsnPheCY 960
|
Db 2869 GAGGAATTCAAAACACGACACGGGTGGCAGAAAGCTGTGGTGTGTTCAAACTTCTTG 2928
QY 960 sLeuPhePheTYrLYsSerH|sGlnAspAsnH|sProLeuAlaSerLeuProLeuLeuG1 980
|
Db 2929 CCTGTCTTCTCAAAATCAACACAGGACAATATCCCTTGGCCAGGCTGTGTGCTCGG 2988
QY 980 YTYrSerLeuThrL|eProSerGlnuSerGlnuAsnL|eGlnLYsAspTYrValPheLYsLe 1000
|
Db 2989 CTATCGCTCAACATCCCTTGTAGTCCAGAACATCCAGAAAGATTAACGTGTTCAGACT 3048
QY 1000 uH|sPheLYsSerH|sValTYrTYrPheArgAlaGlnuSerGluTYrThrPheGlnuArgT 1020
|
Db 3049 GCACTTCAAGTCCCAAGTCTTACTTGAAGGGGGAAGAGAGATTAACGTTTCAAAAGGTG 3108
QY 1020 pMetGlnValL|eArgSerAlaThrSerSerAlaSerArgProH|sValLeuSerH|sLY 1040
|
Db 3109 GATGGAAGATGATCCGAGTGGCAACGAGTCTGTCGCAACCCCAAGTGTGAGCCACAA 3168
QY 1040 sGlnuSerLeuValTYr 1045
|
Db 3169 AGAGTCTTGTGTAT 3184

RESULT 4

Alignment Scores:		
Pred. No.:	2,966-273	length: 3997
Score:	3941.50	Matches: 580
Percent Similarity:	70.0%	Conservative: 162
Best Local Similarity:	54.7%	Mismatches: 261
Query Match:	53.8%	Indels: 17
DB:	7	Gaps: 14
US-09-555-342B-2 (1-1045) x US-10-172-118-1536 (1-3997)		

177 ArgGluHisLeuAlaLysAsnLysTyrTrpProGlnGlnAspAlaLeuGluAspLysIle 136
Db CGAGGACACCTTCAAAGTGAACGATATTTGGCCCTGGCCACAGACACTGCTTGTGAAGAATA 717
197 ValGluPheHisIleAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeu 216
Db CTAGAAATTCATCAGAAGCACTGGGCCAGACACTGCTGAGGTGGATTTCCAGAGTCTC 777
217 GluIleAlaArgArgLeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGlu 236
Db GAAATTCCTCGAAATGGTGAATGTACGGCAGTCAGATTTCACATGGCTTCTGACAGGAA 837
237 GlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
Db GGAAACCAAGATTCACTGGCAGATTCCCAATGGGTGATCTGGTGTTCAGAGGACACACC 897
257 LysIleAsnAlaPheAsnThrAlaLysValArgLysLeuSerPheLysArgLysArgPhe 276
Db AAAATCAACACTTTCACCTGGTCCAAAGTCCGTAATCTAACTTCAAGAGAAAGATTT 957
277 LeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
Db CTTATCAAACTTCATCCAGAGGTTATGACCTTACAGACACATTAAGATTTTGTG 1017
297 AlaSerArgAspPheCysLysSerPheThrLysIleCysValGluHisIleAlaPhePhe 316
Db GGATATTAAGATGAATGTATGAAGACTTCGGAAGATTGTGTGATGATATACACCTTTT 1077
317 ArgLeuPheGluGluProLysProLysProValLeuPheSerArgLysSer 336
Db AGACTTTGGACACACTTAAGCCAAAGAAAGAAAGCGCTTCTTCAGCCGGGCTCTCC 1137
337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyGlnHis 356
Db TTTCAGATTCAGCGGAAGAACTCAGAAACACTAGATTAATTTCAAAACACATGGAATG 1197
357 LysLysValGlnPheGluArgLysHisSerLysIleHis--SerIleArgSerLeu-- 374
Db AAGAAATTCATTAATGAAGAAGGACACGAAAGCCCAACCGTCCGTTCGAGCTTGACT 1257
375 AlaSerGlnProThrGluLeuAsnSerGluValLeuGluGlnSerGlnSerThrSer 394
Db CGAAGCTTACCAAAACAG-----AGC 1278
395 LeuThrPheGlyGlyGlyAlaGluSerProGlyGlyGlnSerCysArgArg----- 411
Db ATCTCATTTCCCCGAGGAGATTGAGACTCTGCTGCCCTCCCATCTTCAGAGCATGCTTTTAC 1338
412 -----GlyLysGluProLysValSerIleArgLysGluProGly----- 423
Db TCCTCTCTCCCTCCACTCTGTGTTCCCTCTGGCTGGCCAGAGTTTAAGACAGACAGC 1398
424 -----SerHisProSerProAlaProArgAspSerProAlaGlyAsnLysGlnAlaAsp 441
Db TCCCTCAAGATCCCCAGAGTTTCCTAACGTCAAGAGTCCAGCTGCAGAGAGGGCGAATGGA 1458
442 GlyAlaAlaSerAlaProThrGlnGluGlnGluGluValValLysAspArgThrGlnGln 461
Db GCAGTGGCTGAGAGGCCCGAC-----ACACATCG 1488
462 SerLysAspProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
Db GCCCAGCCCTTCGGGCCCGCCCGACCTCAGCTTGCTTCAGAGGCTTCCAGAAAGATCCT 1548
479 HisLeuSerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeu 498
Db CAGCCTTCTCCCTCCAGCCGGAAGAGCCGCCGAGCTGAGCGCTGCA-----TTT 1599
499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuLeuAsn 518
Db CAGGTGCTTTGGGCCCTGAGCAAGGAGCTCATCCCACTCTGAGCCCTGCTCCAGT 1659


```

Qy 519 AspGln1aCysProArgThrAspAspGluAspGluYArgArgLysArgPheProThr 538
    |||
Db 1660 GATCTCGTGGGAGCGGGAGTGAATGAGCGAG---GAGGCCAGACACAAGCGGCTCTGGCA 1716
Qy 539 AspLysAlaTyrPheLeu1aLysGluValSerThrThrGluArgThrTyrLeuLysAsp 558
    |||
Db 1717 GACAGAGCTTAACTTCAATGTCMAAGAGATTCTCGCTACAGAACCAATACCTCAAGAT 1776
Qy 559 LeuGluVal1IleThrSerThrPheGlnSerThrValSerLysGluAspAlaMetProGlu 578
    |||
Db 1777 TTAAAGATTATTACCGTGTGTCTTCCGACGCCAGTGTGAAGAGAGCCATCTCTGGC 1836
Qy 579 AlaLeuLysSerLeu1IlePheProAspPheGluProLeuLysPhe1aThrAspPhe 598
    |||
Db 1837 ACTGTGATGACGCGTCTTCTTCCACATGATCCCATCTGTATGATTCCACAGAGGCTTC 1896
Qy 599 LeuLysGlu1IleGluGlnArgLeu1aLeuTyrGluYArgSerAsn1aGln1IleArg 618
    |||
Db 1897 CTGGCGCAGGTGGAGCAGAGGCTGCACTTGGGAAAGGCGCTCCAAAGCCACACAAA 1956
Qy 619 ---AspTyrGlnArg1IleGluAspValMetLeuLysAsn1IleGlnLysMetLys1aLeu 637
    |||
Db 1957 GGCAGTCATCAACGAATCGGGGACATCTCTGCTCAGAAACATGCCCACTTAAAGAGATT 2016
Qy 638 Ala1aHisLeuTyrLys1aSerGlu1aLeuGlu1aLeuGluAsnGly1IleLysSer 657
    |||
Db 2017 ACCAGCACTATTCCAAAGACATGACAGGTCTTAACAGACTGGAAGGCTACCAACGCC 2076
Qy 658 SerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuPro 677
    |||
Db 2077 TGTAAAGAGTTGGAGGACGTGTACAGAGATTGAGCTGCAAGAGGTCTGCTACTTCCCT 2136
Qy 678 LeuAsnThrPheLeuLeuArgProLeuHisArgLysMetHisTyrLysGlnValLeuGlu 697
    |||
Db 2137 CTCAACACGTTCTCTGTAAGCCCATCCAGCGGCTGCTGCACTACCGCTGCTCTGCC 2196
Qy 698 ArgLeuCysLysHis1aProProSerHis1aLysPheArgAspCysArgAla1aLeu 717
    |||
Db 2197 CGCCTTAGCGGACATTAACAGCCCGGACCATGACATACGCTGATGATGAGCCCTG 2256
Qy 718 AlaGlu1IleThrGluMetVal1aGlnLeuHisGlyThrMet1IleLysMetGluAsnPhe 737
    |||
Db 2257 AAAGCCATCAACAGGTGACACACACATTAACAGCACTTCCATCCGCTGGAGAACCTG 2316
Qy 738 GlnLysLeuHisGluLysLysAspLeu1IleGly1IleAspAsnLeuValProGly 757
    |||
Db 2317 CAGAAAGCTAACGAGCTGACAGCGGACCTGTGGACATGAGAACTCATTTGCTCTGGC 2376
Qy 758 ArgGluPhe1aLeuArgLysSerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArg 777
    |||
Db 2377 AGGAGATTCAATCCGTGAGGGCTGCTTCAACAGCTCAACAAAGAGGGCTGACAGAGG 2436
Qy 778 MetPhePheLeuPheAsnAspValLeuLeuTyrThrSerArgLysLeuThrAlaSerAsn 797
    |||
Db 2437 ATGTTTTCCTGTTCTCAGATATGTGTCTGACCAAGCAAGAGAGTTCAGAGGACAGC 2496
Qy 798 GlnPheLysValHisGlyGlnLeuProLeuTyrGlyMetThr1IleGlnLysSerGluAsp 817
    |||
Db 2497 CACTTCGGAGTCCGGGCGCTCTTCCCTCCAAAGCAGTGGCTGGGAAAGAGATPAC 2556
Qy 818 GluTyrGlyValProHisCysLeuThrLeuArgGlyGlnArgGlnSer1aIleVal1a 837
    |||
Db 2557 GAGTGTCTGTTCACACTGTTCACATCAACCGGCTCAAGAAACATGTGTGGCA 2616
Qy 838 AlaSerArgSerArgLysMetGluLysTyrVal1aLysAsp1aIleAspLeu 857
    |||
Db 2617 GCCAGCACTCGCGTGAAGAAAGAGATGCTGACCTGAAGCTCCGCGATCCAAGCA 2676
Qy 858 AlaGluLysSerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
    |||
Db 2677 GCCAAGAGTGGCGGTGACACGCGCCCTGCACTGCCAGGCGCACTGTGTGCACTGTCTCC 2736
Qy 874 ProAspAsnLysSerProAspGlu1aThrAla1aAspGlnLysSerGluAspLeu 893

```

```

Db 2737 CCC-----AGATCCCCCAAGAG---GATCTCTGGAGACAGAGTCAGAAATGATGCT 2787
Qy 894 SerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgLysAsnThrMetValHis 913
    |||
Db 2788 CGGGGTGTCCGAGCTCTCTGAGGGGATGCGCAGCACCAGGCGCAACACATGTCAC 2847
Qy 914 ValCysTyrHisArgAsnThrSerValSerMetValAspPheSer1aLeuValGluAsn 933
    |||
Db 2848 GTGTGTGTGATCCGGAACACACACAGTGTCCAGGGGACAGACACAGTGCAGTGTGAAAC 2907
Qy 934 GlnLeuSerGlyAsnLeuLeuArgLysPheLysAsnSerAsnGlyTyrGlnLysLeuTyr 953
    |||
Db 2908 CAGCTTTCAGAGATATGCTGTMAAGAAAGTTCAAAACAGTCATGCTGGCAGAAAGCTTGG 2967
Qy 954 ValValPheThrAsnPheCysLeuPhePheTyrLysSerHisGlnAspAsnHisProLeu 973
    |||
Db 2968 GTGTCTTTCACAACTTGTGTGTCTTCTTACAAAATCATCATGAGTGAATCCCACTG 3027
Qy 974 AlaSerLeuProLeuLeuGlyTyrSerLeuThr1IleProSerGluSerGluAsn1IleGln 993
    |||
Db 3028 GCCAGCTCTCCGCTGTGCTGGCTACAGGTGAGCATCCCAAGGAGCGCATGGCATACAC 3087
Qy 994 LysAspTyrValPheLysLeuHisPheLysSerHisValYrTyrPheArg1aGluSer 1013
    |||
Db 3088 AAAGACTATGTTTTCAGGCTCCAGTTCAATCCACAGCTCATCTTCCCGGCTGAGAGC 3147
Qy 1014 GluTyrThrPheGluArgTyrMetGluVal1IleArgSer1aThrSerSerAlaSerArg 1033
    |||
Db 3148 AAGTACATTTTAAAGGTGATGAGGTATTCAGGGGCGCAGACCTCAGCCGGAGG 3207

```

RESULT 5

```

US-10-342-887-1536
: Sequence 1536, Application US/10342887
: Publication No. US20040058340A1
: GENERAL INFORMATION:
: APPLICANT: Dai, Hongyue
: APPLICANT: He, Yudong
: APPLICANT: Linsley, Peter S.
: APPLICANT: Mao, Mao
: APPLICANT: Roberts, Christopher J.
: APPLICANT: Van 't Veer, Laura Johanna
: APPLICANT: Van de Vijver, Marc J.
: APPLICANT: Bernard, Rene
: TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
: FILE REFERENCE: 9301-188-999
: CURRENT FILING DATE: 2003-01-15
: PRIOR APPLICATION NUMBER: US/10/342,887
: FILING DATE: 2001-06-18
: PRIOR APPLICATION NUMBER: 60/298,918
: PRIOR FILING DATE: 2002-05-14
: PRIOR APPLICATION NUMBER: 10/172,118
: PRIOR FILING DATE: 2002-06-14
: NUMBER OF SEQ ID NOS: 2699
: SEQ ID NO 1536
: LENGTH: 3997
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-342-887-1536

```

Alignment Scores:

```

Pred. No.: 2,96e-273 Length: 3997
Score: 2941.50 Matches: 580
Percent Similarity: 70.0% Conservative: 162
Best Local Similarity: 54.7% Mismatches: 261
Query Match: 53.8% Indels: 17
DB: 8 Gaps: 14

```

```

US-09-555-342b-2 (1-1045) x US-10-342-887-1536 (1-3997)

```

```

Qy 1 MetGlyGlu1IleGlnGlnArg-----ProThrProGlySerArgLeuGlyAla 16
|||||

```

118 ATGGGGAGATGAGAGACATACAGAGCTCGTGCAGAGCTGCGGTGGTCC 177
Qy 17 ProGluAenSerGlyIleSerThrLeuGluArgGlyGlnIlyProProProThrProSer 36
Db 178 CAGACCCCTGTGGAGTTCAGACCTTGGAGCTGGGCGAGCTCTTGGCCAGATGCA 237
Qy 37 GlyIlyLeuValSerIleValIleGlnMetLeuAspThrGlnGluAlaPheGluVal 56
Db 238 GAGAGGACCTGCACCTCAGATTAAGCTGCTGCAGACACCAAGATTTTGACATT 297
Qy 57 ProGlnArgAlaProGlyIlyValLeuLeuAspAlaValCysAsnHisIleLeuAsnLeuVal 76
Db 298 GAGCTTAATGCGATGGCCAGATTAATTCACACAGTGTGGAAGCCGTTTAACCTGGTA 357
Qy 77 GluGluAspGlyThrPheGlyLeuGluPheProAspHisIlyValIleThrValTyrLeuAsp 96
Db 358 GAAATGTGACTCTTCGGGATGAGATTCMAAATACCTGACCTTGAATTTGGCTTGA 417
Qy 97 LeuLeuIlyAspProIleValIleGlnIleArgArgProIlyValIleValIlyAspPheVal 116
Db 418 CTAATGAAACCAATCATTTAGGCAATACAGAGGCCAAAGATGTGGTCTTCGCTGACT 477
Qy 117 ValIlyAspPheProProAspHisIleThrGlnLeuGlnGluLeuThrArgTyrLeuPhe 136
Db 478 GTAATAATTTTCCACCTGATCCTGGTCAAGCTACAAAGAAATATACAAAGATATCTGTTT 537
Qy 137 AlaIleuGlnValIlyGlnAspLeuAlaGlnIlyArgIleuThrCysAsnAspThrSerAla 156
Db 538 GCCCTTGCAACTTAAAGAGACCTGCTGGAGAGCGTTTGACCTGTGCAGACACACAGCG 597
Qy 157 AlaLeuLeuIleSerHisIleValIleGlnSerGluIleGlyAspPheAspGlnAlaIleuAsp 176
Db 598 GCCCTTCTCAGCTCCCATCTCTCGAGTGGAAATAGAGATTCAGATGAAACGCTGGAC 657
Qy 177 ArgGlnHisLeuAlaIlyAsnIlyThrIleProGlnGlnAspAlaIleuGluAspIlyIle 196
Db 658 GGAAGGACCTCAAGAGTGAACGAGTATTTGCTGGCCAGACAGACCTGCTTGAGAGATA 717
Qy 197 ValGluPheHisIleAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuLeu 216
Db 718 CTAAGATTTCCATCAGAGACAGTGGGCGACACCTGCTGAGTCCGATTTCCAGGTGCTC 777
Qy 217 GluIleAlaArgAlaGluGluMetTyrGlyIleArgLeuHisIleProAlaIlyAspAspGlu 236
Db 778 GAAATGTGCTCAAGATTTGAAATGATGCGCATCAGATTTCAATGCTTGCAGAGGAA 837
Qy 237 GlyThrIlyValLeuAlaIleValAlaAsnThrGlyIleLeuValIlePheGlnIlyPheThr 256
Db 838 GGAACCAAGATTTCAACTGGCAGTTTCCCAATGGGTGATCTGTTTCCAGGGCACCC 897
Qy 257 IlyValIleAsnAlaPheAsnThrAlaIlyValIlyArgIlyLeuSerPheIlyValArgIlyPhe 276
Db 898 AAAATCAACCTTCAACTGCTCCAAAGGTCGTAACCTTCAAGAGGAAAAAGATTT 957
Qy 277 LeuIleIlyLeuAlaGluProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
Db 958 CTAATCAAACTTCAATCCAGAGTTCATGACCTTAACCAAGACATTAAGATTTTGTGTT 1017
Qy 297 AlaSerArgAspPheCysIlySerPheTyrIlyIleCysValIleGlnHisIleAlaPhePhe 316
Db 1018 GGTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
Qy 317 ArgLeuPheGluGluProIlyAspProIlyAspProIlyLeuPheSerArgIlySerSer 336
Db 1078 AGACTTTTGGACCAACCTAAGCCAAAGCCGCTTCTTCAAGCCGGGGCTCCCTCC 1137
Qy 337 PheArgPheSerArgIlyArgThrGlnIlyGlnIleValLeuAspTyrValIlyGluGlyIlyHis 356
Db 1138 TTCAGATCAGTGGAGAGATCTCAGAAACATCAGAGATTTATTTCAAGACAGTGGATG 1197
Qy 357 IlyValIyValIleGlnPheGluArgIlyHisIleSerIlyValIleHisIleSerIleArgSerLeu 374
Db 1198 AAGAGATTCATATGAAAGAGAGACAGCAAGACCCACAGTCCGTTCCAGCTCTGACT 1257

Qy 375 AlaSerGlnProThrGluLeuAsnSerGluValLeuGlnIleSerGlnIleThrSer 394
Db 1258 GCAGACCTTCCAAACACG-----AGC 1278
Qy 395 LeuThrPheGlyGluGluAlaGluSerProGlyGlyIleSerCysArgArg----- 411
Db 1279 ATTCATTTCCCGAGGATTTGAGGACTCTGCTCCATCTTCCAGGAATGCTTTTAC 1338
Qy 412 -----GlyIlyGluProIlyValSerAlaGlyIlyProGly----- 423
Db 1339 TCGCTCTCTCCCTCAGCTGTGTCCTCTGGCTGGCCAGGTTTAAGACAGACAGC 1398
Qy 424 -----SerHisAspSerProAlaProArgArgSerProAlaGlyAsnIlyGlnAlaAsp 441
Db 1399 TCCCTCACAGATCCCGAGGTTTCTTCAAGTCAAGTCCAGCTGCAGAGAGCGAGTGA 1458
Qy 442 GlyAlaAlaSerAlaProThrGlnGluGluIlyValIlyAspArgThrGlnGln 461
Db 1459 GCAGTGGCTGAGAGCCCGAC-----ACACCATCG 1488
Qy 462 SerIlyAspGlnProPro-----GlnProSerThrGlySerLeuThrIlySerPro 478
Db 1489 GCCCAGCCCTCGGGCCCCCGACCTCCAGCTGTGCAGGCTTTCACAGAAAGTCT 1548
Qy 479 HisIleuSerGluLeuSerValAsnSerGlnIlyGlyValAlaProAlaAsnValThrLeu 498
Db 1549 CAGCTTCTCTCCCTCACCCGAGAGAGCCCTCGAGTCAAGCTTGA-----TTT 1599
Qy 499 SerProAsnLeuSerProAspThrIlyGlnAlaSerProLeuIleSerProLeuAsn 518
Db 1600 CAGGTGCTTTGGGCCAGCTGAAACAGGCTCATCCCATCCCGAGCCGCTGCTCAGT 1659
Qy 519 AspGlnAlaCysProArgThrAspAspGluAspGluIlyArgIlyValArgPheProThr 538
Db 1660 GATCTGCGGAGCGGAGCGGAGATGACTGCGAG-----GAGCCAGACACAAAGCGCTGCGCA 1716
Qy 539 AspIlyAlaTyrPheIleAlaIlyGlnValSerThrGlnIlyArgThrTyrLeuIlyAsp 558
Db 1717 GACGAGGCTTATCTTCAATGATGCAAGAGATTTCTCGTCAAGAACCAATACCTTAAGAT 1776
Qy 559 LeuGluValIleThrSerTyrPheGlnSerThrValSerIlyGlnAspAlaMetProGlu 578
Db 1777 TTAGAATTTATACCGTGTGTTCCGACGCCAGTGTGAAGAGACCATCTCGG 1836
Qy 579 AlaLeuIlySerIleIlePheProAsnPheGluProLeuHisIlyValPheHisIleAsnPhe 598
Db 1837 ACTGTGATGACGCTGCTCTTCCAAATGATCCATCTATGAGTTCCACAGAGGCTTC 1896
Qy 599 LeuIlyGluIleGluGlnArgLeuAlaLeuTyrGlnIlyArgSerAlaIleGlnIleArg 618
Db 1897 CTGGCAGAGTGGAGAGAGGCTGGCACTGGGAAAGGCCCTTCAAAAGCCACACAA 1956
Qy 619 ---AspTyrGlnArgIleGlyAspValIleMetLeuIlyAsnIleGlnIlyMetIlyHisIleu 637
Db 1957 GGCAGTATCAACGAATCGGGGACATCTCTGACGAACATGCCCACTTAAGAGATTT 2016
Qy 638 AlaAlaHisIleuTyrIlyHisIleSerGluAlaLeuGlnAlaLeuGlnIlyIleIlySer 657
Db 2017 ACCAGTACTTCCAAAGACATGACGAGTCTTAACAGAACTGAAAGGCTACCAACGC 2076
Qy 658 SerArgArgLeuGlnIlyAsnPheCysArgAspPheGluIlyGlnIlyValCysTyrLeuPro 677
Db 2077 TGTAAAGATTTGAGGAGTGTACAGAGAGTTTGAGTGCAGAAAGTCTGTACTGTGCT 2136
Qy 678 LeuAsnThrPheLeuLeuAspProLeuHisIleArgLeuMetHisIlyTyrGlnValIleGln 697
Db 2137 CTCACACGTTTCCGCTGAGGCCATCAGGCGTGTGCTGACATACCGCTGCTGCGC 2196
Qy 698 ArgLeuCysIlyValHisIleProProSerHisIleAspPheArgAspCysValAlaIleu 717
Db 2197 CGCCTTAGCGGACATTTACAGCCCGGAGCACCATGATACGTGCTGACAGCCCTG 2256

Qy	718	AlaGluIleThrGluMetValAlaGlnLeuHisIleGlyThrMetIleIleMetGlnAspPhe	737
Db	2257	AAAGCCATCAGAGGATGACACACACATCTACAGCACATTTCTATCCGGTGGAGAACTTG	2316
Qy	738	GlnIleLeuHisIleGlnLeuIleValAspLeuIleGlyIleAspAsnIleValAlaProGly	757
Db	2317	CAGAGCTTAACGAGCTGACGAGCGGAGCTGGTGGGATGAGAACTCTATTCCTCTGGC	2376
Qy	758	ArgGlnPheIleArgLeuIleSerLeuSerIleValSerGlyIleGlyIleGlnIleArg	777
Db	2377	AGGAGAGTCAATCCGTAGAGGAGCTGCTTCAACAGCTCAACAGAAAGGCGCTGCAGAGAGG	2436
Qy	778	MetPhePheLeuPheAsnAspValLeuLeuIleIleValIleThrSerArgGlyLeuThrAlaSerAsn	797
Db	2437	ATGTTTTCCTGCTTCCAGATATGTTGCTGTACACAGAAAGAGTTGCAGGAGCACAGC	2496
Qy	798	GlnPheIleValHisIleGlyIleLeuProLeuIleIleGlyMetThrIleGlnIleSerGlnAsp	817
Db	2497	CACCTCCGGATCCCGGGGCTCTCTCCCTCCCAAGGATGCTGGTGAAGAAAGTGAATTAAC	2556
Qy	818	GluTrpGlyValProHisCysLeuThrIleAsnIleGlyIleGlnIleSerIleIleValAla	837
Db	2557	GAGTGGTCTGTTCCACATCTGTTCCACCATCTACGCGGCTCAGAAACAACTGGTGGCA	2616
Qy	838	AlaSerSerArgSerGluMetGluIleIleGlyIleTrpValGluAspIleGlnMetAlaIleAspLeu	857
Db	2617	GCCAGACATCTCGGTGAGAAAGAGAGTGGATGTCAGACTGAATCCGCGCATCAAGCA	2676
Qy	858	AlaGlnIleSerSerSerProAlaProGlnPhe-----LeuAlaSerSerPro	873
Db	2677	GCCAAAGATGGCCGGTACACGCGCCCTGCATGCGACGCGGCGCATGCTGTGCCCTCTGCC	2736
Qy	874	ProAspAsnIleSerProAspArgIleAlaThrAlaAlaAspGlnIleSerGluAspAspLeu	893
Db	2737	CCC-----AGATCCCCCAACAG---GATCTCTGAGACAGAGATCAAGAAATGATGCT	2787
Qy	894	SerAlaSerArgThrSerLeuIleArgIleAlaProHisArgGlyAsnThrMetValHis	913
Db	2788	CGGGGTGTCGCGACGCTCCCTGGAGGGGCAATGCGACACCGGGCCAAACACCAATGAC	2847
Qy	914	ValCysTrpHisArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn	933
Db	2848	GTGTGCTGTGATCCGGAACACACGCTGTCCAGGCAAGACACACAGTGCATGTGCAGAAC	2907
Qy	934	GlnLeuSerGlyAsnLeuLeuArgIleValSerAsnSerAsnGlyIleTrpGlnIleLeuTrp	953
Db	2908	CACCTTTCAGAGATATGCTTAAGAAAGTTCAAAAACAGTCAATGCTGGCAAAAGCTCTGG	2967
Qy	954	ValValPheThrAsnPheCysLeuPhePheTrpIleValSerHisGlnAspAsnHisProLeu	973
Db	2968	GTGTGCTTACCAACTCTGTTGTTCTTCTACAAACCTCAACAGATGATCACTCCACTG	3027
Qy	974	AlaSerLeuProLeuLeuGlyIleTrpSerLeuThrIleProSerGluSerGluAsnIleGln	993
Db	3028	GCCAGCTCCCGGCTGCTGGGCTACACAGCGACATCCCAAGGAGGCGCGATGACATAC	3087
Qy	994	IleAspTrpValPheIleValHisPheIleValTrpIleValTrpPheArgAlaGluSer	1013
Db	3088	AAAGACTATAGTTTCAAGCTCCAGTTCAATCCACAGTCAATCTTTCGCGGCTGAGAGC	3147
Qy	1014	GluTrpThrPheGluArgTrpMetGluValIleArgSerAlaThrSerSerAlaSerArg	1033
Db	3148	AACTTACATTTGAAGAGTGGATGAGAGTGATCCAGGGGGCCAGAGCTGACCGGGAGG	3207
RESULT 6			
US-10-802-432-16			
; Sequence 16, Application US/10802432			
; Publication No. US20040185489A1			
; GENERAL INFORMATION:			
; APPLICANT: Hoffmann-La Roche Inc.			
; TITLE OF INVENTION: Transcriptional Activity Assay			
; FILE REFERENCE: 21574			
; CURRENT APPLICATION NUMBER: US/10/802, 432			

```

CURRENT FILING DATE: 2004-03-17
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn version 3.1
SEQ ID NO 16
LENGTH: 3997
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: KIAA0793
LOCATION: (1)..(3997)
OTHER INFORMATION: GenBank accession No. AB018336
US-10-802-432-16

Alignment Scores:
Pred. No.: 2,966-273 Length: 3997
Score: 2941.50 Matches: 560
Percent Similarity: 70.0% Conservative: 162
Best Local Similarity: 54.7% Mismatches: 261
Query Match: 53.8% Indels: 57
DB: 9 Gaps: 14

US-09-555-342B-2 (1-1045) x US-10-802-432-16 (1-3997)

QY 1 MetGlyGluIleGluGlnArg-----ProThrProGlySerArgLeuGlyAla 16
Db 118 ATGGGGGAGATGAAAGAAACATACAGAGTCTGCAGACATGCAGGGAGTGGCGTGGGCGCC 177
QY 17 ProGluAnseSerGlyIleSerThrLeuGluArgGlyGlnLysProProThrProSer 36
Db 178 CAGACCCCTGGGGAGTTTACACCTTGAGCCTGGGACAGCTCTCTGGCCAGAAATGCAA 237
QY 37 GlyLysLeuValSerIleLysIleGlnMetLeuAspSerThrGlnGluAlaPheGluVal 56
Db 238 GAGAGGACACTGCACCTCAGAGTTAAAGCTGCTGCACACACATGGAAATATTGGACATT 297
QY 57 ProGlnArgAlaProGlyLysValIleuLeuAspAlaValCysAsnHisIleuLeuVal 76
Db 298 GAGCTTAATAGCGATGGCCAGGATATTACTGACACAAAGTGTGAGAGCGTTTAACTGTGTA 357
QY 77 GlnGlyAspTyrPheGlyLeuGlnPheProAspHisIleLysValIleThrValTrpLeuAsp 96
Db 358 GAAATGTGACTACTTGGGATGGAGTTTCAAATATACAGTCTGACTGATTTGGATTGGCTTGA 417
QY 97 LeuLeuLysProIleValIleGlnIleArgGlyProLysHisIleValValIleLysPheVal 116
Db 418 CCTAAGAAACCCATCATTAAGGAAATACGAAAGCCCAAAAGATGTGTGCTTGCCCTAGCT 477
QY 117 ValLysPhePheProProAspHisIleThrGlnLeuGlnGlnGluLeuThrArgTyrLeuPhe 136
Db 478 GTAAATTTTTTCCACTGATCTCTGGTCAAGTCAAGAAAGAAATATACAAAGATCTTGTT 537
QY 137 AlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspHisSerAla 156
Db 538 GCGTTTGCAACTTAAAGAGAGACCTGCTGGAAAGACGCTTTGACCTGTGCTTCAACACACGCG 597
QY 157 AlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAspGlnAlaLeuAsp 176
Db 598 GCGCTTTCACGCTCCCACTCTCTGCGAGTTCGAAATTTGGAGATTACGATGAACGCTGGAC 657
QY 177 ArgGlnHisIleuAlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGlnLysPyrIle 196
Db 658 CGAGAGCACCTCAAGATGAAAGAGATTTGCTGTGGCCAGACAGACAGCTGCTTGAGAGATA 717
QY 197 ValGlnPheHisIleAsnHisIleGlyGlnThrProAlaGlnUserAspPheGlnLeuLeu 216
Db 718 CTAGAATTTCCATCAGAAAGACCTGGGACAGACACCTGCTGAGTGGGATTTTCCAGGCTTC 777
QY 217 GlnIleAlaArgArgLeuGlnUwerTyrGlyIleArgLeuHisProAlaLysAspArgGlu 236
Db 778 GAAATGTGCTCGAAATTTGGAAATTTGACGGCAATCAAGATTTCAACATGCTTCTGACAGGAA 837
QY 237 GlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256

```

D	b	898	GGGAACCAAGATTCAATGGCACTTTCCACATGGGTGTACTGTGTCCAGGGCACAC	897
O	y	257	LYSII LEAMNLA PHEAENITPRLALYVAIARGLYLEUSERPHELYAHTGLYARPH	276
D	b	898	AAAATCAACACTTCAACTGATGCTCAAGGCTCGTTAACTMACTCAAGGAAAAGATT	957
O	y	277	LEUILELYLEUAHXPROMAPRLAENSERIALATYRGINAEPTHTLEUGLUPHEUMET	296
D	b	958	CTTATCAAACTTCATCCAGAGGTTATGACCTTACAGGACACATYTGAAATTTTGTTG	1017
O	y	297	ALASERARGPHECYELYSERPHETPLYSILECYVAIGLHIHIALAPHEP	316
D	b	1018	GGTATTAAGATGAAATGTAAGAACTTCTGGAAGATTGTGTGAGATACACCTTTT	1077
O	y	317	ARGLEUPHEGLUGLUPROLYBPROLYBPROLYALEUPHESEARGLYSESER	336
D	b	1078	AGACTTTTGGACCAACCTAAGCCAAAGAAAAGCGCTCTTCTCAGCGGGGCTCTCC	1133
O	y	337	PHEAHTPHESESGLYARGTHRGINLYSGINVAILEAERTYRVALLYGLUGLYHIS	356
D	b	1138	TYCAATATCAAGTGAAGAACTCAGAAACAACTAGTAGATTATTTCAAAGACAGTGGAA	1197
O	y	357	LYSLYVAIAGINPHEGLIARGLYHISERLYSILEHIS---SECTIARSERLEU---	374
D	b	1198	AAAGAAATTTCCATTATAGAAAGAAAGGACACCAAGACCCACACGTCGTTGACTGACT	1255
O	y	375	ALASERGINPROTHRGILEUAMNSERGINVALILEUGINSESGINSETHSER	394
D	b	1258	GCAGACCTACCAAAACAG-----AGC 1278	
O	y	395	LEUTHRPHIEGLYGLIYALAGLUSERPROGLYGLINSECYAHTARG-----	411
D	b	1279	ATCTCATTTCCCGAAGGAAATTGAGACCTCGCTCCCATCTTCACGAAATGCTTTAC	1338
O	y	412	-----GLYBGLUPROLYVALAMSERLAGLYGLUPROGLY-----	423
D	b	1339	TGCGTCTCTCCCTCCACTGTGTCTCCCTGTGCTGCCAGAGTTTAAGACACAGCAGC	1398
O	y	424	-----SERHAPROSERPROALAPROARGASERPROALAGLYAMLYSGINLAAP	441
D	b	1399	TCCCTCAAGATCCCAAGTTTCTCAAGTCAGAAATCCAGCTGCAGAAAGCGCGATGA	1455
O	y	442	GLYALALASERIALAPROTHRGILUGLUGLUGLUGLUGLUGLUGLUGLUGLUGL	461
D	b	1459	GCAGTGTGAGAGGCCCGAC-----ACACCARG 1488	
O	y	462	SERYAPROGINPRO-----GLNPROSETHRGLYSERLEUTHGLYSEPRO	478
D	b	1489	GCCCCAGCCCCCTCGGGCCCCCGCAGCTCCAGCTGTGTCCAGGCTTTCCACAAAGTCTT	1548
O	y	479	HISLEUSERGILUSERSERVALAMSERGINGLYVALALAPROALAMVALTHLEU	498
D	b	1549	CAGCTTTTCTCCTCAAGCCGGAAGAGCCCCCTGAAGTGTGAGCCCTGCA-----TTT	1599
O	y	499	SERPROAMLEUSERPROAPRTHRYSGINLASERPROLEUILESERPROLEUAMN	518
D	b	1600	CAGGTGCTTTGGGGCCAGCTGMAAAGAGGTCATCCCACTCTGAGCCGTGTCTAGT	1659
O	y	519	APRGINALACYBPROARGTHRAPRARGILUAPRGILYATGATGLYAHTPHEPROTH	538
D	b	1660	GATGTGTGCGAGAGCCGGAGTGAATGCTGAG--GAGCCACAGACAAACGGGTGCTTCA	1716
O	y	539	APRLYVALATYRPHIEALALYSGILUALSERTHTRHGRGLUARGTHTYLEULYAPR	558
D	b	1717	GACGAGGCTTACTTATATGTCMAAAGATTTCTCGCTACAGAAACGAACTTACAGAT	1776
O	y	559	LEUGLUALIETHSERTRIPHEGLINSETHRVASERLYSGIUAAPALAMEPROGLU	578
D	b	1777	TTAGAAAGTTATTACGTGTGTTGCTCGACGCGAGTGTGTAAGAGAGAGCGCATGCTCG	1836
O	y	579	ALALEULYSERLIEPHEPROANPHEGLUPROLEUHIALLYBPHHITHRANPHE	598
D	b	1837	ACTGTGAAGAGCTGTCTTCTCCAAATCGATCCCATCTTATGAGTTTCAACAGAGGCTTC	1896

QY	559	LeuLySgUllleGlUGlaRgLeuLaleuTPrgUllYARgSerAsnAlglInLeArg	618
Db	1897	CTGGCGAGGTGAGACAGAGGGCTGGGACTCTGGAAAGGGCCCTCCAAAGCCACACAAA	1955
QY	619	---AaPYrGlnAArgIleGlyAsrValMeLLeuLyAsnAlleGlInGlyMeLysHleu	637
Db	1957	GGCAGATCATCAACGATGGGGGACATCTCGTCAGAGAACATGGCGCAATTAAAGAGTTT	2011
QY	638	AlaAlaHleLeuTPryHHisSerGluLaleuGlUalaleuGlUasngUlylleYsSer	657
Db	2017	ACCACCTACTCTCCAAAGACATACACAGGCTCTTAAAGAACTGGAAAGGCTACCAAGCGC	2076
QY	658	SeTARgARgLeuGluLAsnPhrCyARgAspPhrGluLeuGlnLyValCySTryLeuPro	677
Db	2077	TGTAAAGAAAGTTGGAGGCACTGACAGAGATTGAGCTGCAAGAAAGTCTGCTACTTGGCT	2133
QY	678	LeuAenThrPhrLeuLeuLargProLeuHlsARgLeuMeHlsrTyLySgInValleuGlu	697
Db	2137	CTCAACACGTTCTCTGTAAGCCCATCCAGCGGCTGTGCACTACCGCTGTGTCGCGC	2199
QY	698	ArgLeuCyLyVHHisrProProSerHlsAlaAspPhrAspCyARgAlaLaleu	717
Db	2197	CGCCTATGGGACATTAACAGCCCGGGGACCATGACTAGCTAGCTGCGCATAGACCCCTCG	2255
QY	718	AlaGluIleThrGluMeValAlaGlnleuHlsGlyThrMeLleLyMeGluAsnPhr	737
Db	2257	AAAGGCATCACAGAGGTATACCAACACATACAGACATTTCTATCGGCTGGAGAACTCG	2311
QY	738	GlnLyLeuHlsGlnLeuLyLeAspLeuIleGlylleAspAsnLeuValAlaProGly	757
Db	2317	CAGAAAGCTAAACGAGACTCAGCGGGGACTGTGGGGCAATAGAGAACTCATGCTCTCGGC	2376
QY	758	ArgGluPhrIleArgLeuGlySerLeuSerLyLsAsrGlyLyGlyLeuGlnGlnArg	777
Db	2377	AGGAGTTCAATCCGTAGAGGCTGCTTCAACAGCTCACAGAAAGGCTTGACAGAGAG	2433
QY	778	MetPhrPhrLeuPhrAsnAspValleuLeuTyThrSerArgGlyLeuThrAlaSerAsn	797
Db	2437	ATGTTTTTCTGTTCTCAGATATGTTGCTGTACAAAGCAAAAGAGTTGCAGAGCACAGC	2496
QY	798	GlnPhrLyValHlsGlyGlnLeuProLeuTyrgLyMeThrIleGlnGlnSerGluAsp	817
Db	2497	CAC TTCGGAGATCCGGGGCTCTCTCCCTCCAAAGGACATGCTGGTGAAGAAAGTATATAC	2555
QY	818	GluTPrgLyAlaProHlsCyLeuThrLeuArgGlyGlnArgGlnSerlleValAla	837
Db	2557	GAGTGTCTGTCCACACTGTTCACACATCAACGGGCTCAAGAAACAATCTGTGTGCA	2611
QY	838	AlaSerSerArgSerGluMeGluTyTrpValGluAspIleGlnMeAlaIleAspLeu	857
Db	2617	GCCAGACATCCGGCTGGAGAAAGAAAGATGTGCTGACCTGAACTCCGCATCCAAAGCA	2676
QY	858	AlaGluLySserSerSerProAlaProGluHe-----LeuAlaSerSerPro	873
Db	2677	GCCAAAGAGTGGCGGTGACACGGCCCTCTCACTGCCAGGCGCGCACTGTGTGCACTGTCCC	2733
QY	874	ProAspAenLySserProAspGlyAlaThrAlaAlaAspGlnGlnSerGluAspAspLeu	893
Db	2737	CCC-----AGATCCCCCAAGAG---GTATCTGTGAGCAGAGATCGAAGATGATGCT	2789
QY	894	SerAlaSerArgThrSerLeuGlnArgGlnAlaProHlsARgGlyAsnThrMeValHls	913
Db	2788	CGGGGTGTCGACAGCTCCCTGAAGGGGAGTGGACAGACCGGGCCAAACACCAATGAC	2847
QY	914	ValCySTrPHlsARgAenThrSerValSerMeValaAspPhrSerlleAlaValGluAsn	933
Db	2848	GTGTGTCTGTACCGGAACACAGCTGTCCAGGGAGACCAACATGACGCTGTGAGAAC	2907
QY	934	GlnLeuSerGlyAsnLeuLeuArgLyPhrLyAsnSerAsnGlyTPrgInLyLeuTP	953
Db	2908	CAGCTTTTCAGATATCTGCTAAAGAAAGTTCAAAACAGTCAATGGCTGGCAAGAGCTCGG	2967


```

Oy      1025 ArgSerAlaThrSerSerAlaSerArgProHisValIleuSerHisArgGluSerIleuVal 1044
Db      1442 CGCGATGCGACCGACGCTGCGCTCGCGACCCCAAGGTGGAGCCACCAAGAGCTCTTGTG 1501
Oy      1045 Tyr 1045
          |||
Db      1502 TAT 1504

RESULT 8
US-09-764-868-51
/ Sequence 51, Application US/09764868
/ Patent No. US20020168711A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OR INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PT232
/ CURRENT APPLICATION NUMBER: US/09/764.868
/ CURRENT FILING DATE: 2001-01-17
/ Prior application data removed - refer to PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 1510
/ SOFTWARE: PatentIn Ver. 2.00
/ SEQ ID NO 51
/ LENGTH: 1718
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-764-868-51

```

Alignment Scores:			
Pred. No.:	2.6e-86	Length:	1718
Score:	1006.50	Matches:	189
Percent Similarity:	78.7%	Conservative:	47
Best Local Similarity:	63.0%	Mismatches:	57
Query Match:	18.4%	Indels:	7
DB:	3	Gaps:	3
US-09-555-342B-2 (1-1045) x US-09-764-868-51 (1-1718)			
QY	738	GlnlyvLeuHhlgLnuLeuLybAerLeuIlegIYlEAaRvbnLeuValProGly	757
Db	13	CAGAAgCTAAcGAGAGCTCGAGCGGGAGCCTGAGGGCATAGAGAAcCTCATGTCTCGGC	72
QY	758	ArgGluPheIleArgLeuGlySerLeuSerLybLeuSerGlylybGlyLeuGlnGlnArg	777
Db	73	AGGAGTTCATCCGTGAGGGCTGCCCTTCAcAAAGCTCAcAAAGAAAGGCGCTGCAGcAGAGG	132
QY	778	MetPhePheLeuPheAaRvbnArValLeuLeuYrThSerArgGlyLeuThrAlaSerAsn	797
Db	133	ATGTTTTTCTCTTCTCAATATATGTTGCTGTACACAAcGAAAGAGTTGCAGGAGCAGC	192
QY	798	GlnPheLybValHhlgLyGlnLeuProLeuYrGlyMetThrIleGluGlnSerGluAsp	817
Db	193	CACTTCCGGATCCGGGGCTCTCTCCCTCCAAAGGATCTGTGTGAGAAAGTGAATAC	252
QY	818	GluTrpGlyValProHhGcybLeuThrLeuAaRgLyGlnAaRgInserIleIleValAla	857
Db	253	GAGTGTCTGTTCcCAACTGTTTCAcCATCTACCTACCGGCTCCAGAAACAAATCGTGTGGCA	312
QY	838	AlaSerSerArgSerGlnMetGlnYrSTrPValGluAspIleGlnMetAlaIleAspLeu	857
Db	313	GCCAGACATCCGGCTGGAGAAAGAAAGAACTGGATCTGACCTGAATCTCCGGATCCAAAGCA	372
QY	858	AlaGluLybSerSerSerProAlaProGluPhe-----LeuAlaSerSerPro	873
Db	373	GCCAAAGAGTGGCGGTGACACGGCGCCCTTGACCTGCCAGGCGCGCACTGTGTGACCTCGTCC	432
QY	874	ProAaRvbnLybSerProAaRgIleAlaThrAlaAlaAaRgInGlnSerGluAaRvbnLeu	893
Db	433	CCC-----AGATCCCCCAACAG---GATATCTGGAGAGGAGAGTCAAGAAATGATGCT	483
QY	894	SerAlaSerArgYrThSerLeuGluArgGlnAlaProHhIaRgLybAnThrMetValHhS	913
Db	484	CGGGGTGTCCGAGACTCTCTGGAGGGGCAATGGCCACACCGGGGCTCAACACACAAATGGAC	543

[illegible]

```

RESULT 9
US-11-097-143-3041
; Sequence 3041, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: C1000728
; CURRENT FILING DATE: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157, 832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160, 191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161, 932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164, 769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173, 383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175, 693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184, 831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191, 637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3041
; LENGTH: 1747
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-3041

Alignment Scores:
Pred. No.:      2,756-77      Length:      1747
Score:          913.50        Matches:      217
Percent Similarity: 54.0%      Conservative: 110
Best Local Similarity: 35.8%    Mismatches:   207
Query Match:    16.7%         Indels:       73
DB:             13            Gaps:         14

US-09-555-342B-2 (1-1045) x US-11-097-143-3041 (1-1747)

```


; PRIOR APPLICATION NUMBER: 60/175,693
 ; PRIOR FILING DATE: 2000-01-12
 ; PRIOR APPLICATION NUMBER: 60/184,831
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/191,637
 ; PRIOR FILING DATE: 2000-03-23
 ; NUMBER OF SEQ ID NOS: 43008
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 743
 ; LENGTH: 1220
 ; TYPE: DNA
 ; ORGANISM: DROSOPHILA
 US-11-097-143-743

Alignment Scores:
 Pred. No.: 9.4e-77 Length: 1220
 Score: 905.50 Matches: 180
 Percent Similarity: 67.84 Conservative: 69
 Best Local Similarity: 49.04 Mismatches: 105
 Query Match: 16.64 Indels: 13
 DB: 13 Gaps: 7

US-09-555-342B-2 (1-1045) x US-11-097-143-743 (1-1220)

QY 8 ProThrProGlySerArgLeuGlyValAlaProGluAaSerGlyIleSerThrLeuGluArg 27
 DB 102 CCAAGGGGGGGGGCTGCTGCGGAGGCGGCATGACGCACTCCCTCAACACTCTTCGCGG 161
 QY 28 GlyAluProProProProProProProProProProProProProProProProProPro 44
 DB 162 GTGATGACGAGCCCTCCACACCCGACATCGCGGCGGCAAAAAGCTGACTGTTCGATC 221
 QY 45 GlnMetLeuAaPhePheThrGlnGluAlaPheGluValAlaProGluAaPheGlyIleVal 64
 DB 222 CAGATGCTGAGCAGCTGATCCATCCATGTTCCAAAGTACAGGCTTAAAGCACTGGCGCGTG 281
 QY 65 LeuLeuAaPheValAlaCysAaPheValAlaGluGlyAaPheGlyIleValAlaPheGlyIle 84
 DB 282 CTGTTTGAACAGATGATGCGGCTGATGATTTACTGAGGCGGCACTTCTGCGCTTGAG 341
 QY 85 PheProAaPheProProProProProProProProProProProProProProProProPro 101
 DB 342 TACAGGAAGTCTCCACATACCAAA-----TACTGGCTGATCTGGAAGACCAATG 395
 QY 102 ValLeuGlnIleAaGArgProLeu---HisValAlaValAlaPheValAlaPhePhe 120
 DB 396 AACGCCAGGCGGCTCTCTTATTCATCCAGTCCGCTTCTGATCAAGTTCTAT 455
 QY 121 ProProAaPheIleThrGlnLeuGlnGluGluLeuThrArgTyrLeuPheAlaLeuGlnVal 140
 DB 456 ACACCCGATCCAGGCAATTGGAAGAGATACACCAAGTATTTCTTCTGCGCAATC 515
 QY 141 LysGlnAaPheValAlaGlnGlyAaGlyLeuThrCysAaAaPheThrSerAlaAlaLeuLeuIle 160
 DB 516 AACGAGACCTGGCCACAGGCAAGTCTGCAAGTCAAGTAAACACGCGGCTTATGCA 575
 QY 161 SerHisIleValAlaGlnSerGluIleGlyAaPhePhe-----AaPheValAlaLeuAaPheGlu 178
 DB 576 AGCTAATTTTGAACGCGCTCATGCGGCACTTCGTCGAGGAGTATTCACGACATC 635
 QY 179 HisLeuAlaLysAaPheValAlaPheProGlnGlnAaPheAla-----LeuGluAaPheValAla 197
 DB 636 TATCTGCTCTGATCGCTTGTGCTGCGCAACGAGCGCACCATCATCAGCGAAGTATTCG 695
 QY 198 GluPheHisIleAaPheHisIleGlyAlaThrProAlaGluSerAaPheGlnLeuLeuGlu 217
 DB 696 GAGAAACCAAAACATGTTGTCATCCCGGCGGAGGCGGACCTTAACCTTGGAG 755
 QY 218 IleAlaAaGArgLeuGluMetCysTyrIleAaGluHisIleProAlaLysAaPheGlyIle 237
 DB 756 ACGCTCGAAGTGTGAGCTGATGCAAGAAAGCATCCGCGCAAGAGTGTGGAAGG 815
 QY 238 ThrIleIleAaPheValAlaAlaAaPheThrGlyIleLeuValAlaPheGlnGlyPheThrLys 257

DB 816 GTGCGCTTATCTGCTGTGCGCCACATGGGCATCACAGCTTCCADAACATCAGCGG 875
 QY 258 IleAaPheAaPheThrAlaLysValAlaGlyLeuSerPheIleValAlaPhePheLeu 277
 DB 876 ATCAACACCTTCTCGTGGCTTAAGTACGAAGATTTCTTCAAGCGCAAGGATTCCTG 935
 QY 278 IleLysLeuAaGArgProAaPheAlaAaSerAlaTyrGlnAaPheThrLeuGluPheLeuVal 297
 DB 936 GTCAACTGATCCGAGAGATGATGATTAACAGATACCGCTGAGTTCTTCTTCGAG 995
 QY 298 SerAaPhePheCysLysSerPheThrPheIleCysValAlaGluHisIleAlaPhePheArg 317
 DB 996 GGTGCAACGAGTGCAGAAACCTTGGAAAAATCGTCGAAATCAGGATTCCTCGA 1055
 QY 318 LeuPheGlu---GluProLysProLysProLysProValAlaPheSerThrGlySer 336
 DB 1056 TGCACCTGCCGTAATAACGCCACGCGCAAAATCGCTTCTCGCGGGGTAGTTCA 1115
 QY 337 PheArgPheSerGlyArgThrGlnLeuGlnValLeuAaPheGlyValAlaGluGlyHis 356
 DB 1116 TTCGCTATGAGGGAAGAAACCAAGAGCATTAATGAGTTCTGCGGAAATTAATGTCG 1175
 QY 357 LysLysValAlaGlnPheGluArg 363
 DB 1176 AACGCCCAAACTTCCAAAG 1196

RESULT 11

; Sequence 3040, Application US/11097143
 ; Publication No. US20050208558A1
 ; GENERAL INFORMATION:

; APPLICANT: Ventec, J. Craig

; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
 ; FILE REFERENCE: CLO00728

; CURRENT APPLICATION NUMBER: US/11/097,143

; PRIOR FILING DATE: 2005-04-04

; PRIOR APPLICATION NUMBER: 60/157,832

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: 60/160,191

; PRIOR FILING DATE: 1999-10-19

; PRIOR APPLICATION NUMBER: 60/161,932

; PRIOR FILING DATE: 1999-10-28

; PRIOR APPLICATION NUMBER: 60/164,769

; PRIOR FILING DATE: 1999-11-12

; PRIOR APPLICATION NUMBER: 60/173,383

; PRIOR FILING DATE: 1999-12-28

; PRIOR APPLICATION NUMBER: 60/175,693

; PRIOR FILING DATE: 2000-01-12

; PRIOR APPLICATION NUMBER: 60/184,831

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: 60/191,637

; PRIOR FILING DATE: 2000-03-23

; NUMBER OF SEQ ID NOS: 43008

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3040

; LENGTH: 4150

; TYPE: DNA

; ORGANISM: DROSOPHILA

US-11-097-143-3040

Alignment Scores:
 Pred. No.: 1.42e-74 Length: 4150
 Score: 891.50 Matches: 222
 Percent Similarity: 45.4% Conservative: 113
 Best Local Similarity: 30.1% Mismatches: 201
 Query Match: 16.3% Indels: 202
 DB: 13 Gaps: 15

US-09-555-342B-2 (1-1045) x US-11-097-143-3040 (1-4150)

QY	423	GIYserHiIvProSerProAlaProArgSerProAlaGIYvniLyvSglnAlaSerLy	442
DB	3064	GGATCGGTAAGTCTTCCGTCACTATCCGATCTTGAGGAGGAACA-----	3020
QY	443	AlaIaSerAlaProThGlnGlnGlnGlnGlnValValLyvAspArgThGlnGlnSer	462
DB	3019	-----TATGACGTAAATACGGCAAAATCCAGACAGCT	2987
QY	463	LyvProGlnProGlnProSerThGlySerLeuThGlySerProHiLeuSerGln	482
DB	2986	AAGAGACAG-----GAGCTAGCCACAGCGCGTGGCACGCGCTCG-----	2948
QY	483	LeuSerValaenSerGlnGlyGlyValaIaProAlaen--ValThLeuSerProAsn	501
DB	2947	-----CACACAAATGGGAACGGAAACGGAAATGGGCATATCCCTAGC-----	2906
QY	502	LeuSerProAspThrLyvGlnAlaSerProLeuIleSerProLeuLeuAspGlnAla	521
DB	2906	-----2906	
QY	522	CysProArgThrAspAspGlnAspGlnGlyArgArgLyvArgPheProThAspLyvAla	541
DB	2905	---ACACAAACGATATGAGGCGAGAGATGAAGAAACGGAAATGGCCACCGAGCCGAC	2849
QY	542	TyrPheIleAlaLyvSglnValSerThrThGlnValGthTrTyLyvLyvAspLeuGlnVal	561
DB	2848	TACTTTTGGCCGAGAGACTGTGTGAATGACGAAACCTAGCTACAGAAAGATCTGATGTG	2789
QY	562	IleThrSerTrpPheGlnSerThrValSerLyvSglnAspAlaMetProGlnAlaLeuLyv	581
DB	2788	CTGAACACAACTTCCGACAGGTGTGAGTCTCGGGAGTGTG-----GACACAGTCCAG	2735
QY	582	SerLeuIlePheProAsnPheGlnProLeuHiLyvSerPheHiSerThAsnPheLyvSgln	601
DB	2734	CCGCTG---TTCGAGCTCTGCACTCCCTGGCGCCAGCACCAATCTCTTCTGGCGCAC	2678
QY	602	IleGlnGlnAlaArgLeuAlaLeuTrGlnGlnLyvArgSerAsnAlaGlnIleArgAspTrpGln	621
DB	2677	ATCGAACAACCCGATGTGTGACGTGGAGGAGGAAGAGAGC-----CACGAAGCCCAAC	2627
QY	622	ArgIleGlyAspValMetLeuLyvAsnIleGlnGlyMetClyvHiLyvAlaAlaHiLyv	641
DB	2626	CGCATCGAGACGTATATGATGAAGACATGGCCGCTGTGCCCATTCAGATGATGAGTGTG	2567
QY	642	TrpLyvHiSerGlnAlaLeuGlnAlaLeuGlnAsnGlyIleLyvSerSerArgArgLeu	661
DB	2566	CAGACGCACTCGACATCTTGCACTGTATGAACGACATGTACGAAGGCGATGAACGCTTC	2507
QY	662	GlnAsnPheCysValArgAspPheGlnLeuGlnLyvValCysTrpLyvProLeuAsnThrPhe	681
DB	2506	CGTCAAGGTGTATCAAGAGTTTGTAGCGAAGAAAGTTTGTCTATCCATCGGCGAACTT	2447
QY	682	LeuLeuArgProLeuHiValArgLeuMetHiSerTrpLyvSglnValIleGlnAlaArgCysLyv	701
DB	2446	CTACTGAAGCCCTTCAACGCTGTGTGAATTCACACAGATCTTAAAGCGGCTGTGGAC	2387
QY	702	HiSerProProSerHiValaAspPheArgAspCysValGlnAlaLeuAlaGlnIleThr	721
DB	2386	TACTATGGGAGAGACATATGACATACCCGATGTATGAGCCGATGACCACTTGCTGCT	2327
QY	722	GlnMetValAlaGlnLeuHiGlyLyvThMetIleLyvMetGlnAsnPheGlnLyvLeuHi	741
DB	2326	CGCAGACCAAGGATATTAAGTCCGAGCTTCCGACTGTGCAAACTTTGTGAGACGTGT	2267
QY	742	GlnLeuLyvLyvAspLeuIleGlyIleAspAsnLeuValIaProGlyArgGlnPheIle	761
DB	2266	GAACTCGCAACCGGAC---ATCAACTTCGACAGATTTGTTCAGCCCACTGCGGCTATTC	2210
QY	762	ArgLeuGlySerLeuSerLyvLeuSerGlyLyvGlyLeuGlnGlnIaMetPhePheLeu	781
DB	2209	CGCCACGGGATGCTTTCTGAAGACACTCCAAAGCGCGGTGTGCAGCAGAGGATTTCTTCTTG	2150

QY 762 PheaaNpPValleuLeuTYrThrSerArgIylLeuThrAlaSeranGlnPheLysVal 801
Db 2149 TTTCTCCGACCTGCTCTCTACCGTTTGGAAATCC---CCGCTGGACCAAGAGCTTCGGCATT 2093
QY HisGIyGlnLeuProLeuTYrGIyMetThrIleGlnGluSerGluNpGluTYrGlyVal 821
Db 2092 TTAGGCCACAGTTCCTGCTGCGTTTCATTACTACCCGAGAACGCCGAGAC--- 2045
QY 822 ProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAspSerArg 841
Db 2044 ---AACACCTTCTCCATCTTCGGCGGACAGTGTGCATCACAAGTGAATGGCGGACACC 1988
QY 842 SerGluMetGluLYrValGluNpIle---GlnMetAlaIleAspLeu----- 857
Db 1987 GCTGAAGAAAGACCTCTGCTGGCGGACGCTTCGAAAGCGCGCGGATATTAAAAACCGA 1328
QY 857 ----- 857
Db 1927 CCGCCCAACATGCAACTCCAGCTAACGACGTCAGAAATCGAGTGAATATATATTG 1868
QY 857 ----- 857
Db 1867 TATAAATGCGATTGCGTAGGTTTGTTCATATAGACGTTTAAACAAAGTCCATATA 1808
QY 857 ----- 857
Db 1807 TGAATTGAATAGCTTTAAAGAACCAAGCTACCGGAAATATATTAAATATCAAGTAAATCTT 1748
QY 857 ----- 857
Db 1747 ATAAACCAAAAGCTTGTGTCGACGCGCTAAACCTATATACATCAACAAAGTAAAAAGCTTT 1688
QY 857 ----- 857
Db 1687 AACTTAAACGTACCTTTGTTAAACTCCCTGTTTCTCCATTGTGTTTGTGCTGGC 1628
QY 858 -----AlaGluLYrSerSerSerProAlaProGluNpLeu 869
Db 1627 AAGAAAGACAGATCATGACGCGCGGTAAAGCAAGAGATGCACCCAAATCCCAAAATCTT 1568
QY 869 ----- 869
Db 1567 CTTCGAAAGACCCAGTAAATGTTATTTCACCTTTATTTCAGGCTCTCGAGAGAGG 1508
QY 870 -----AlaSerSerProProAspNpLYrSerProAspGluAlaThrAla 885
Db 1507 CCTGACCTATTTCGCTCTGAGCAAGGAAACAAACAGACCTGAAACAGCAGTGTAAACG 1448
QY 886 AspGlnGluSerGlu-----AspAspLeuSerAlaSerArgThrSerLeuGluArg 902
Db 1447 CGGCGGTGCGCTGACTTAACGACAGCAGCAAGATGACAGCTGCGACGACACACGCAAAACAG 1388
QY 903 GlnAla-ProHisArgGlyAsnThrMetValHisValCysTrpHisArgAsnThrSerVal 922
Db 1387 GACGAGGCGCATCCCGAAGCAATACCGGCTCTGCACGCTCTTTGGCATCTGGCGCCACCGT 1328
QY 922 IserMetValAspPheSerIleAlaValGluNpGlnLeuSerGlyAsnLeuLeuArgly 942
Db 1327 AGGACATGGGCGATCACTGATATAGCCGCGGACACCAAGTATGTCGGGTATATCTGTCGAAA 1268
QY 942 sPheLYrAsnSerAsnGlyTYrPGLnLYrLeuTYrValValPheThrAsnPheCysLeuNp 962
Db 1267 GTTCAAGAAATAGCTCCGCTGGCAAGACTCTGGGGTGGTTCACGCTTTTGTCTGTA 1208
QY 962 sPheTYrLYrSerHisGlnAspAsnHisProLeuAlaSerLeuProLeuLeuGlyTYrSe 982
Db 1207 CTTTACCAAGAGCTACCAAGACGAAATTCGCACTGGCTATCTGCGCATTTGGGTTACAC 1148
QY 982 rLeuThrIleProSerGluSerGluAsnIleGlnLYrAspTYrValPheLYrLeuHis sPn 1002
Db 1147 GGTGGGTCTCTCTGTCTACCAAGAGCCGTTTCAGAAAGAGTTTCGCTTCAAGCTTCTT 1088
QY 1002 eLYrSerHisValTYrTYrPheArgAlaGluSerGluTYrTYrThrPheGluArg 1019

Db 1087 TAAAGACACGCTACTTCTTCCGGCGGAAAGTCGACACCTTACAAAG 1036
RESULT 12
US-10-755-889-499
; Sequence 499, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; PRIOR FILING DATE: 2004-01-13
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 499
; LENGTH: 3309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-499

Alignment Scores:
Pred. No.: 7,2e-61 Length: 3309
Score: 748.50 Matches: 195
Percent Similarity: 49.6% Conservative: 94
Best Local Similarity: 33.4% Mismatches: 233
Query Match: 13.7% Indels: 61
Gaps: 15

US-09-555-342b-2 (1-1045) x US-10-755-889-499 (1-3309)
Qy 18 GluSenSerGlyLLeSerThrLeuGluArgGly-----GlnLysProPro 32
Db 331 CAGAAATATCTAGACGTAACCTCTCTCGGTCCATTAAAGATTGTCAAAAGCT--- 387
Qy 33 ProThrProSerGlyLysLeuValSerLLeuValLeuMetLeuAspThrGlnGlu 52
Db 388 -----AAAGCATGCAGTGCAGAAAGTGAATCTTCTGATGAATCAGAAATAT 432
Qy 53 AlaPheGluValProGlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHis 72
Db 433 ACTGTGTATGAGAAACGCTCCAGAGACAAAGTGTGATTAAGTGTGTAACAC 492
Qy 73 LeuMetLeuValGluGlyAspTyrPheGlyLeuGluPheProAspHisLysLeuValThr 92
Db 493 TTGAATCTGTAGAGAAAGACTACTTGGCTTACGTATCCAGATGCTGAACACAGAG 552
Qy 93 ValTyrLeuAspLeuLysProLLeuValLysGlnLLeuArgArgProLys---HisVal 111
Db 553 AATGTGTGACCCCTGTAGAGAAATTAATAAAACAGGTTCCAAAGTGTGTCGAC--- 609
Qy 112 ValValLysPheValValLysPhePheProPheAspHisThrGlnLeuGlnGluLeu 131
Db 610 ---TTTCATTATTATGTAAATTATATCCACCAAGCCTGCCCAACTATCTGAAGATATC 666
Qy 132 ThrArgTyrLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCys 151
Db 667 ACCAGGATCTACCTGTCTGTGAGTGGAGATGACATGCTGTCCGAAAGGCTGCCCTGC 726
Qy 152 AsnAspThrSerAlaAlaLeuLeuLLeuSerHisLLeuValGlnSerGlnLLeuValPhe 171
Db 727 TCTTTGTTACCTGCTGCTGTGGCTCTCACTGACCTGTCCAGTCAAGACTCGAGACTAT 786
Qy 172 Asp---GluAlaLeuAspArgGlnHisLysLeuValLysLeuTyrLLeuProGlnLIn--- 189
Db 787 GACCCAGATGATGTGGAGCGATTACATAGTAGTTCGCTTTGCAACCAACCAACT 846
Qy 190 AspAlaLeuGluAspLysLLeuValGluPheHisLLeuAsnHisLLeuGlyGlnThrProAla 209

Db 847 AAAGACTGGAAGACAAAGTATGAGCTGACCAAGACCAAGAGAAATGACCGCACCA 906
Qy 210 GluSerAspPheGlnLeuGlnLLeuAlaArgArgLeuGlnMetTyrGlyLLeuArgLeu 229
Db 907 GAAGCAGAGATGCAATTTCTGGAATAATGCCAAAATTTATCAATGTATGGGTGGATTTA 966
Qy 230 HisProAlaLysAspArgGluGlyThrLysLLeuMetAlaValAlaAsnThrGlyLLe 249
Db 967 CATCATGCTTAAGACTCAGAAAGGGGTAGAAATTAATGTAAGAGTTTGTGCAAGTGTCTG 1026
Qy 250 LeuValPheGlnGlyPheThrLysLLeuMetAlaPheAsnThrAlaLysValArgLysLeu 269
Db 1027 TTGATATATGCGACCCGCTGCGAATAAAGATTTGCTGCGCCCAAGCTTCAAAATTT 1086
Qy 270 SerPheLysValGlyLysArgPheLeuLLeuLysLeuArgProAspAlaAsnSerAlaTyrGln 289
Db 1087 TCATTAACAAACGGAACAACTTTTACATTAAGATCCGCGCGAGAGTTTGAACAAATTTGA 1146
Qy 290 AspThrLeuGluPheLeuMetAlaSerArgAspPheCysLysSerPheThrLysLLeuCys 309
Db 1147 AGCAGCATGGGATTTAAGCTGCCAAACCATGACGCTGCCAAAGCTTTATGAAATGATATG 1206
Qy 310 ValGlnHisHisAlaPhePheArgLeuPheGlnGluProLysProLysProLysProVal 329
Db 1207 GTTGAAGCATCATACATTTTTCAGACTACTG---TTACCAAGAACACCTCCCAAGAAATTC 1263
Qy 330 LeuPheSerArgGlySerSerPheArgPheSerGlyValGlnThrGlnLysGlnValLeuAsp 349
Db 1264 CTA---ACCTTGGGTTCACAAATTTCTGTAATGATGTCAGAGAACACAAACGAGAAAGA 1320
Qy 350 TyrValLysGluGlyLysLLeuValGlnPheGluArgLysLLeuSerLysLLeuHis 369
Db 1321 GCCAGCTGCTGATTAAGATGCCSCACACCTTACTTGAACCTATCCAGCAAACTGAT 1380
Qy 370 SerLLeu---ArgSerLeuAlaSerGlnProThrGlnLeuAsnSerGluVal---LeuGlu 387
Db 1381 ACATGATGTCGACAGTTGATGAGACATGACATGAAATGAAACATGAATATATACATGAAG 1440
Qy 388 GlnSerGlnGlnSerThrSerLeuThrPheGlyGlnGlyAlaGlnSerProGly----- 405
Db 1441 GATTCTATGTCGTGCGACAGATTTGTACTGCGCAGTACGCGCACAAACAAAGGCACTCT 1500
Qy 406 -----GlyLysSerCysArgArgGly 412
Db 1501 CAGACCAACTGATACCACTGTGACTCCGAGAAAGAGCTGAGAGAGAGCGGAGCAGAG 1560
Qy 413 LysGluProLysValSerAlaGlyLysProGlySerHisProSerProAlaProArgArg 432
Db 1561 GAAGAGACAAACGAGAGAGGGGAGAGAAATGACACGCCCATCTGCGCATCCAGACAGAG 1620
Qy 433 SerProAlaGlyAsnLysGlnAlaAspArgLLeuAlaAsnSerAlaProThrGlnGlnGlu 452
Db 1621 GGAAGACTGACAGTGAACGCGACGACACCGCCAGCGGAGAGACCACTGCCACTAG 1680
Qy 453 GluValValLysAspArgThrGlnGlnSerLysProGlnProProGlnProSerThrGly 472
Db 1681 GAGCTAGAAAATACTCAAGATGACCTGATGAACATCA----- 1719
Qy 473 SerLeuThrLysSerProHisLLeuSerGluLeuSerValAsnSerGlnGlyValAla 492
Db 1720 -----ACCAACATTTAGAGACTGAAGAAACCTTTTGAACACTCAACAGACACTGCC 1773
Qy 493 ProAlaAsn-----ValThrLysSerProAsnLeuSerProAspThrLys 507
Db 1774 GTAACGATGATGAGAGAGAGAGAGCTTTCCACTCCCGCTGCGACATGCGCGCCAGGAG 1833
Qy 508 GlnAlaSerProLeuLLeuSerProLeuLeuAsnArgGlnAlaCysProArgThrAspAsp 527
Db 1834 GAGAGTGCCTCCATGATGAACCACTTGTCCCTGAAGAGAAATG---GAAACCAAGACG 1890
Qy 528 GluAspArgGlyArgArgPhePheThrAspLysAlaTyrPheLLeuLysGlu 547

Db 1891 GAGTCACGTGATAGACGAGAACCCACCG----- 1920
Qy 548 ValSerThrThrGluArgThrTyrLeuLysAspLeuGluValIleThrSerTyrPheGln 567
Db 1921 ---TGCACCCACCTCCGCTTAGCACTGAGAGGTGGTCAGAGAACCGTGTGTGGAGG 1977
Qy 568 SerThrVal 570
Db 1978 AGCCGGCGTG 1986

RESULT 13
US-10-956-157-1800
; Sequence 1800, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mount, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1800
; LENGTH: 3309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-1800

Alignment Scores:
Pred. No.: 7.2e-61 Length: 3309
Score: 748.50 Matches: 195
Percent Similarity: 49.6% Conservative: 94
Best Local Similarity: 33.4% Mismatches: 233
Query Match: 10.7% Indels: 61
Gaps: 15

US-09-555-342B-2 (1-1045) x US-10-956-157-1800 (1-3309)
Qy 18 GluAsnSerGlyIleSerThrLeuGluArgGly-----GlnLysProPro 32
Db 331 CAGAAATCATCTAGACAGTAAACTCTCGGTCTCATTAAGATTGCAGAAAACCT--- 387
Qy 33 ProThrProSerGlyLysLeuValSerIleValIleGlnMetLeuAspThrGlnGlu 52
Db 388 -----AAAGCATGCAGTGCAGAGTGAATCTTCGATGCATGCAGAAATAT 432
Qy 53 AlaPheGluValProGlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHis 72
Db 433 ACCGTGATGTAGAGAAACGCTCCAGAGCAAGTGGCTTGTATTAAGTGTGAAACAC 492
Qy 73 LeuAsnLeuValGluGlyAspTyrPheGlyLeuGluPheProAspHisLysLysIleThr 92
Db 493 TTGAATCTGTAGAGAAAGACTACTTTGGGCTTACGTATCGATGCTGAAACACAGAG 552
Qy 93 ValTyrLeuAspLeuLysProIleValIleGlnIleArgArgProLys---HisVal 111
Db 553 AATTGGTTGACCTCTGTAAGAAATATAAAACAGGTTCGAAAGTGTCTGGCAC--- 609
Qy 112 ValValLysPheValValLysPheProProAspHisThrGlnLeuGlnGluLeu 131
Db 610 ---TTTTCATTATATGGAATTTTATCCACAGACCTGCCAACTATCTGAAAGATATC 666
Qy 132 ThrArgTyrLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCys 151
Db 667 ACCGAGTACTACCTCTGTCGACATGCGAGATGACATCGTGTCCGAAAGGCTGCCCTGC 726
Qy 152 AsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPhe 171
Db 727 TCTCTTGTATACCTGGCTGTGGCTCCCTACCTGCTCCAGTCCAGTACAGAGCTAT 786
Qy 172 Asp---GluAlaLeuAspArgGlnHisLeuValAlaLysValLysValTyrIleProGlnGln--- 189

Db 787 GACCCAGATGAATGTGGAGAGCATTTACATTAGTAGAGTTCCGCTTGCACCAACCACT 846
Qy 190 AspAlaLeuGluAspLysIleValGluPheHisIleAsnHisIleGlyGlnThrProAla 209
Db 847 AAGAACTGAGAGCAAAAGTATCGAGCTGCACAAAGGCCACAGAGAAATGACCCAGCA 906
Qy 210 GluSerAspPheGlnLeuLeuGluIleAlaArgArgLeuGluMetTyrGlyIleArgLeu 229
Db 907 GAGCAGAGATGCATTTCTTGAAATATCCAAAATTAATCATGTATGGGTGATTTA 966
Qy 230 HisProAlaLysAspArgGluGlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIle 249
Db 967 CATCATGCTAAGAGCTCAGAAAGGGTGAATATATTGTAGAGATTGGCAAGTGGCTGTG 1026
Qy 250 LeuValPheGlnGlyPheThrLysIleAsnAlaPheAsnThrAlaLysValArgLysLeu 269
Db 1027 TTGATTAATTCGCGACCGCGTCGAAATMAAGATTTCCCTGGCCCAAGGTTCTAAAGATT 1086
Qy 270 SerPheLysArgLysArgPheLeuIleLysLeuAspProAspAlaAsnSerAlaTyrGln 289
Db 1087 TCATCAAAACGGAACAATTTTACATTTAAGATCCGCGCGGAGAGTTGAACAATTGAA 1146
Qy 290 AspThrLeuGluPheLeuMetAlaSerArgAspPheCysLysSerPheTyrLysIleCys 309
Db 1147 AGACCATGGGTTTAACTGCCAAACCATCGAGCTGCAGAGCGTTATGGAAGTATGT 1206
Qy 310 ValGlnHisIleAlaPhePheArgLeuPheGluGluProLysProLysProVal 329
Db 1207 GTTGAGCATCATCATTTTTCAGACTACTG---TTACCAAGACCACTCCCAAGAAATTC 1263
Qy 330 LeuPheSerArgLysSerSerPheArgPheSerGlyArgThrGlnLysGlnValLeuAsp 349
Db 1264 CTA---ACCTTGGTTTCAAGTTTCCGTATAGTGGCAGAGCACAGGCCCAAGAGAA 1320
Qy 350 TyrValLysGluGlyGlyHisLysLysValGlnPheGluArgLysHisSerLysIleHis 369
Db 1321 GCCAGTGCATTGATAGATGCCAGCACCTTACTTTAGACGTCATCCAGCAACGTTAT 1380
Qy 370 SerIle---ArgSerLeuAlaSerGlnProThrGluLeuAsnSerGluVal---LeuGlu 387
Db 1381 ACCATGCTCTGCAGCTTGATGAGATCACTGAAATGAAACCATGAAATATACATGAG 1440
Qy 388 GlnSerGlnGlnSerThrSerLeuThrPheGlyGluGlyAlaGlnSerProGly----- 405
Db 1441 GATTCTATGTCTGCTGAGAGGTGGTACTGGCCAGTACGCCCAACAAAGGATCTCT 1500
Qy 406 -----GlyGlnSerCysArgArgGly 412
Db 1501 CAGACCAACTTGATCACCACTGTGACTCCGAGAGAAAGGCTGAGAGCGGAGCAGAG 1560
Qy 413 LysGluProLysValSerIleArgLysIleProGlySerHisIleProSerProAlaProArg 432
Db 1561 GAAAGAGACAAACGAGAGAGGGGAAAGATGACGCCCATCTTCGCGCATTCAGACAGAG 1620
Qy 433 SerProAlaGlyAsnLysGlnAlaAspGlyAlaAlaSerAlaProThrGlnGluGlu 452
Db 1621 GGAAGACTCACTAGACCGCAGCAGACCGCAGACCGGAGAGACCATGCGACAGTGG 1680
Qy 453 GluValValLysAspArgThrGlnGlnSerLysProGlnProProGlnProSerThrGly 472
Db 1681 GAGCTGAAGAAAACCTCAAGATGACCTGATGAACATCA----- 1719
Qy 473 SerLeuThrGlySerProHisLysSerGluLysSerValAsnSerGlnGlyValAla 492
Db 1720 -----ACCAAACTTAGCGAGTGAAGAAACCTTTTAAACCTTCACAGACACTGCC 1773
Qy 493 ProAlaAsn-----ValThrLeuSerProAsnLeuSerProAspThrLys 507
Db 1774 GTAACGAATGAATGAGAGAGAGGCTTTCACCTCCCGTGGAGCTGGCGGCAAGCAG 1833
Qy 508 GlnAlaSerProLeuIleSerProLeuLeuAsnAspGlnAlaCysProArgThrAspAsp 527
Db 527

Db 1894 GAGATGCCCATGATGATGAAACCACTTGCTCCGTGAAGAAATG---GAAACCAAGCG 1890
Qy 528 GUAAGGUGLYAAGARGLYAARGPHEProThraPlybAlaTYrPheIleAlaLYeGlu 547
Db 1891 GAGTCCAGTATGAGACGGAACCCACCG----- 1920
Qy 548 ValSerThrGluArgThrTYrIleuLYsApleuGluValIleThSerTyrPheGln 567
Db 1921 ---TGCACCACTGCGCGCTTGACACTGAAAGGTGTGCAGAGACCGTGTGTGAGG 1977
Qy 568 SerThrVal 570
Db 1978 AGCGGCGTG 1986
RESULT 14
US-10-152-319A-2048
Sequence 2048, Application US/10152319A
Publication No. US20040072160A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Higgs, Brandon
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
TITLE OR INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5089-US
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: US 60/292,335
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/297,523
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,925
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,810
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,807
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,808
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/315,047
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: US 60/324,928
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US 60/330,867
PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: US 60/330,462
PRIOR FILING DATE: 2001-10-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2221
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2048
LENGTH: 4543
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. NM_053927
US-10-152-319A-2048
Alignment Scores:
Pred. No.: 3,216-60 Length: 4543
Score: 744.00 Matches: 318
Percent Similarity: 40.14 Conservative: 159
Best Local Similarity: 26.74 Mismatch: 405
Query Match: 13.64 Indels: 313
DB: 8 Gaps: 41
US-09-555-342B-2 (1-1045) x US-10-152-319A-2048 (1-4543)
Qy 6 GlnArgProThrProGlySerArgLeuGluValAProGluAenSerGlyIleSerThrLeu 25
Db 275 CAGGACAAAGAACCGGAGCTTCGCGCGCGCAGCTGCGAAACAGCTTGATGATCAGACATT 334

Qy 26 -GluArgGlyGlnIlyProProThrProSerGlyIlybLeu----- 39
Db 335 TGAAGACGATTAAGCTCTCTCAGAAAGCTCAGAGCAAACTCTCACGCTCCCATTTGAA 394
Qy 40 -----ValSerIleLYeIleGlnMetLeuAAspThrGlnI 52
Db 395 GATCGTAAAGAGCGCTAAAGACATGCACTGCAAGAGTCAAGCTTGATGATGATGATGAT 454
Qy 52 uAlaPheGluValAProGlnArgAlaProGlyLYsValIleuLeuAAlaLYsAAsnI 72
Db 455 CGCGTGTATGTATGATGAAGCGCTCCGAGGACAGATGCTGTTGACAAAGTGTGCAACA 514
Qy 72 sLeuAAsnLeuValGlnGlyAspTYrPheGlyLeuGluPheProAAspIlybLYsIleTh 92
Db 515 TCTGAACCTGCTTAAGAAAAGACTACTGCTTCACGATGACAGCGAAGATTCAGAA 574
Qy 92 rValTYrLeuAAspLeuLeuLYsProIleValIlyGlnIleArgArgProLYs---HisVa 111
Db 575 GAACGTGTGACCTCGCTAAGAAATTTAAACAGATTGCAAGTGTGCTTGCAAT-- 632
Qy 111 lValValLYsPheValIlybPheProProAAspHisThrGlnLeuGlnGluLe 131
Db 633 ----TTTTCATTTATGTGAAGTTTACCCACAGACCTGCGCATATCAGAAAGATAT 688
Qy 131 uThrArgTYrIleuPheAlaLeuGlnValIlybGlnAAspLeuAlaGlnIlyArgLeuThrCY 151
Db 689 CACCAGGATTAACCTGCTTGACGTGCGAGTGAATGATGTGTCTGACCGCTTACCCTG 748
Qy 151 sAsnAAspThrSerAlaAlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAAspH 171
Db 749 TTCTCTGTGACCTCGCTGCTGCGCTCTTAACAGATGAGTCAAGCTTGCGGACTA 808
Qy 171 eAAsp---GluAlaLeuAAspArgGluHisLeuAlaLYsAAsnLYsTYrIleProGlnGln- 189
Db 809 TGATCTCGATGATGATGGAATGAAATGACTATTAAGAGATTCGCTTGCGCAAAACACAC 868
Qy 190 -AspAlaLeuGluAAspLYsIleValGlnPheHisHisAAsnIleIleGlnIleThrProAl 209
Db 869 GAAAGAACTGAGGATGAAGATGATGAGCTGCACAAAGACACAGAGAAATGACGCCAGC 928
Qy 209 aGluSerAAspPheGlnLeuLeuGluIleAlaArgArgLeuGluMetTYrGlyIleArgLYe 229
Db 929 CGAGGACAGATGACATCTCTCGAGGAATGCCAAAGCTCCCATGTRTGGGGTGAAGATT 988
Qy 229 uHisProAlaLYsAAspArgGlnGlyThrLYsIleAAsnLeuAlaValAAsnThrGlyTI 249
Db 989 ACACCAAGCAAAAGATTCGGAAGGCGTGAAGATCATGTAGGGGTGTGTGCCAGCGGCT 1048
Qy 249 eLeuValPheGlnIlybThrLYsIleAAsnAlaPheAAsnThrAlaLYsValArgLYsLe 269
Db 1049 GTTGATCTATCTGTAACCGGCTTCGAAATTAACAGATTTGCTTGCGCAAGGCTTAAAT 1108
Qy 269 uSerPheLYsArgLYsArgPheLeuIleLYsLeuArgProAAspAlaAAsnSerAlaTYrGI 289
Db 1109 TTCTTATTAACGGAACAACATTTTACATTAAATCCGGCCAGAGATTTGAACAATTCGA 1168
Qy 289 nAAspThrLeuGluPheLeuMetAlaSerArgAAspPheCYsLYsSerPheThrIlybLYeCY 309
Db 1169 AAGCACCATTGGCTTCAAGTTGCGCAACCATAGGGCTGCCAAGCGCTTGGAAGATATG 1228
Qy 309 eValGlnHisHisAlaPhePheArgLeuPheGlnGluInProLYsProLYsProLYsProVa 329
Db 1229 TGTGAAACATCAATCAATTTTCAAGTACTG--CTACCAAGAACACCTCCAAAGAAATT 1285
Qy 329 lLeuPheSerArgLYsSerPheArgPheSerGlyArgThrGlnIlybGlnValLeuAs 349
Db 1286 CCTG---ACCTTGGGCTTCAAGTTCCGCTACAGTGGCAAGAACCCAGGCGGAGAGAG 1342
Qy 349 pTYrValIlybGlnIlybLYsIlybLYsValGlnPheGlnArgLYsHisSerLYsIleHI 369
Db 1343 AGCCAGTGACATTGATAGACCGCGCGGCGCTTACTTCCGAACGCTGCTCCAGCAAGATTA 1402


```

Qy 369 sserlle---ArgSerLeuAlaSerGlnProThrGluLeuAsnSerGluValLeuGluG1 388
Db 1403 CACCATCTTCGCGAGCTTGATGAGGCTCAGTAAGTAAACATGAAATATATC----- 1457
Qy 388 nserGlnGlnSerTherSerLeuThrPheGlyGlyAlaGluSerProGlyGlyGlnSe 408
Db 1457 ----- 1457
Qy 408 rcyArxArgGlyLyagGluProLybValSerAlaGlyGluProGlySer-HisProSerP 428
Db 1458 -----ATGAAGATTCTATGTCCTGTCAGAGGTTGTGTACAGGCCAGTAAGC 1504
Qy 428 roAlaProArGArgSerSerProAlaGlyAsnLySglnAlaAspGlyAlaAlaSerAlaProT 448
Db 1505 CACAACCAAGGTCATCTCT-CAGACCAACCTGATCAC-----ACTGTGACCCCGAGA 1557
Qy 448 hrtGluGluGluGluValValLyAsp-ArgThrGlnGlnSerLybProGlnProPro 467
Db 1558 AAAAGGCTGAGAGAGACACGTGAGAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAG 1617
Qy 468 GlbProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAsnSer 487
Db 1618 CGCCCGTCGAGACCCCTCGGCGACAGAGGAAAGTAC----- 1653
Qy 488 GlnGlyLyValAlaProAlaAsnValThrLeuSerPro-----AsnLeuSerPro 504
Db 1654 ---CAGGGCATGCTCGACATCATGTCCC-CTGTCACTTCATCAGGCCCATCTTGACCT 1709
Qy 505 -----AspThrLySglnAlaSerProLeuIle 513
Db 1710 CCCTCTCCACAGAGCTCCGTAGAGGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1763
Qy 514 SerProLeuLeuAsnAspGlnAlaCysProArGThrAspAspGluAspGluGlyArgArg 533
Db 1764 -----GGTTCTGCTGTGCTTGAGACAAG----- 1787
Qy 534 LybArGpHeProThrAspLybAlaTyR-PheIleAlaLySglnValSerThrThrGluArg 553
Db 1788 -----CTGTCTCGAGGTCTGATGCCAGGGAG 1817
Qy 554 ThrTyLeuLybAspLeuGluValIleThrSerThrPheGln----- 567
Db 1818 GCCCTACCTAGGGGACCAAGAGTGGCATTTAGTACAGGAGCCGCGTGGCAAGAAC 1877
Qy 568 SerThrAlaSerLySglnAspAlaMetProGluAlaLeuLySglnIle----- 584
Db 1878 ACACTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1937
Qy 585 -----PheProAsnPheGluProLeuHisLybPheHisThrAsnPheLeuLyS 600
Db 1938 GGGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1982
Qy 601 GluIleGluGlnAlaLeuThrGluGlyArgSerAsnAlaGlnIleArgAspTyR 620
Db 1983 CAG----- 1985
Qy 621 GlnArgIleGlyAspValMetLeuLybAsnIleGlnGlyMetLybHisLeuAlaAlaHis 640
Db 1985 ----- 1985
Qy 641 LeuTrpLybHisSerGluAlaLeuGluAlaLeuGluAlaGlyIleLybSerSerArgArg 660
Db 1986 AGTTGGCAGAGATTCTTGCC-----ATCCGCTCGCCCTCTCTCTCTCTCTCTCTCT 2024
Qy 661 LeuGluAsnPheCysArgAspPheGluLeuGlnLySglnValCysTyRLeuProLeuAsnThr 680
Db 2025 CTTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2080
Qy 681 PheLeuLeuProLeuHisArgLeuMetHisTyRLySglnVal----- 695
Db 2081 -----ATAGCCCT---CACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2131
Qy 696 -----LeuGluArgLeuCysLyb-HisHisProProSerHisAlaAspPh 710

```

```

Db 2132 GCCAAGCGCGCTCTCTGAGCGCTCCCTAGACATGACCCGAGTGACAGTTTCAGAG-- 2189
Qy 710 eaArgAspCysArgAlaAlaLeuAlaGluIleThrGluMetValAla----- 725
Db 2190 -----GAAAGACTGACGTAGCGAGACAGACACTGCGCGTGTAGAGAGACAG 2239
Qy 726 -----GlnLeuHisGlyThrMetCileLybMetGluAsnPheGlnLybLe 740
Db 2240 TGCCACAGAGCTCGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2299
Qy 740 whis---GluLeuLybLybAspLeuIleGlyIleAspAsnLeuValProGlyArgG1 759
Db 2300 TCAAGATGAGCTGTGAAACACCAACCAATATATACGAATG-----AAAAGAAC 2350
Qy 759 urPheIleArgLeuGlySerLeuSerLybLybSerGlyLybGlyLeuGlnGlnArgMetPh 779
Db 2351 CTCTTGAACAACTCTTACAGAACTGCTTACCAACCAAGATGGAAAGAG----- 2402
Qy 779 ePheLeuPheAsnAspValLeuLeuTyRThrSerArgGlyLeuThrAlaSerAsnGlnPh 799
Db 2403 -----CTTCCACATCCCGTGGCGCTGGCTGCCAGGACGA 2440
Qy 799 elySValHisGlyGlnLeuProLeuTyRgLyMetThrIleGluGluSerGluAspGluTr 819
Db 2441 GATGCGGCCCATGATTGAGCACTGTACTGAGAGAGATTAACAAATCTTCTGGGGAGAA 2500
Qy 819 p-----GlyValProHisCysLeuThrLeuAspGlyGlnArgGlnSerIleIleVa 836
Db 2501 GCTCATGATAGTGTCCGAGATCTCTCAGTCTATTAAGTCTGCGAGAGAACCCACAGATT 2560
Qy 836 lAlaAlaSerSerArgSerGluMetGluLybTyRValGluAspIleGlnMetAlaIleAs 856
Db 2561 CATAGAGGGGGTTTCACTTACTACCCAGAGCTGGTGGCAGAACTGGAAAGAACTGA 2620
Qy 856 pLeuAlaGluLybSerSerSerProAlaPro----- 866
Db 2621 GACCATTTGAACAGAGATGAAACCCACCCACCTCAGCCCTCAGCGAGAGAGGT 2680
Qy 866 ----- 866
Db 2681 GTTTCAGAGAAACGTGCTGTGTAGAGAGAGACGTAAGTGAACGTGACGCCAGTGGGA 2740
Qy 867 -----GluPheLeuAlaSerSerProProAsnLy 877
Db 2741 CGCTTCTCACACAGCCAGGGATGATGTGATGCCACAGATCAGACCCGCAAGCCGTCA 2800
Qy 877 sSerProAsp----- 880
Db 2801 CAGTGGGAATGGAGAGAGGCTTTCTGTGACAGAGACGCTTAAGAAACAGAGAGAGA 2860
Qy 881 -----GluAlaThrAlaAlaAspGlnIle-----SerGluAs 891
Db 2861 GAGAGCTGACAGATCTGCCCCGAGAGAGAAACAACAGACCACTGTCTCCAGAGAGAGA 2920
Qy 891 pAspLeuSerAlaSerArgThrSer-----LeuGluArgGlnAlaProHisArgGlyAs 909
Db 2921 TCAGGTGTGGGCATCCATCTCTCTGAGGGTTGGAACAAAGTCT----- 2966
Qy 909 nThrMetValHisValCysTrpHisArgAsnThrSerValSerMetValAspPheSerI1 929
Db 2967 -----CATTTGAGTGTGCTCACTGTAAGGTAGAA---AGCAT 3001
Qy 929 eaIaValaGluAsnGlnLeuSerGlybAsnLeuAlaGlybPheLybAsnSerAsnGlyTr 949
Db 3002 AAGTGTGGCAGAGTCTCTCCAGAGAGAGTG-----AAGCTAGACCTTAC 3049
Qy 949 pGlnLybLeuTrpValValPheThrAsnPheCysLeuPhePheTyRLybSerHisGlnAs 969
Db 3050 CAGAGAGGTGCCAGTGTCCACAGAGAGAAACATCACTACAGAGCTTCTCAGGT 3109
Qy 969 pAsnHisProLeuAlaSerLeuPro-----LeuLeuGlyTyRSerLeuThrIleProSe 987

```

Db 3110 CGAT---CCTGGGGAGATCTGAGCCAGCGTGTCTATAGCGCCCAAAGATCAGTC 3166
Qy 987 rGluSerGluuunllleglnlybAepTYrValPheIySerHlsValTY 1007
Db 3167 TGAAGACCGAGTACTACACACACACACATATACCAAAACTGAAAGAGCATTC 3226
Qy 1007 rTYrPheArgIaglu-SerGlu-----TyrThrPheg 1018
Db 3227 GGAAGACGAGATTGAGAGCGAATAGTCATCAGGGAGATCCGACATTGACATGACCA 3286
Qy 1018 luaTYrPheGluValle-----ArgSerAlaThrSerSerAlaSerArgProH 1035
Db 3287 GGGCTGGCTCAGGCAATTAAGAGCCCAAGAGAGAGACCCGACATGTCAGTAACCA 3346
Qy 1035 lsvaIleuSerHlslyeGluSer 1042
Db 3347 A-GTAGTGCTCCATTAAGAGACA 3368

RESULT 15
US-11-036-196-2048
Sequence 2048, Application US/11036196
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Johnson, Kory
APPLICANT: Higgs, Brandon
APPLICANT: Casale, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5089-US
CURRENT APPLICATION NUMBER: US/11/036,196
CURRENT FILING DATE: 2005-01-18
PRIOR APPLICATION NUMBER: US/10/152,319
PRIOR FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: US 60/292,335
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/297,523
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,925
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,810
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,807
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,808
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/315,047
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: US 60/324,928
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US 60/330,867
PRIOR FILING DATE: 2001-11-01
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2221
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2048
LENGTH: 4543
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. NM_053927
US-11-036-196-2048

Alignment Scores:
Pred. No.: 3,21e-60 Length: 4543
Score: 744.00 Matches: 318
Percent Similarity: 40.1% Conservative: 159
Best Local Similarity: 26.7% Mismatches: 405
Query Match: 13.6% Indels: 313
Gaps: 41
US-09-555-342B-2 (1-1045) X US-11-036-196-2048 (1-4543)

Qy 6 GluArgProThrProGlySerArgLeuGlyAlaProGluuanserGlyIleSerThrLeu 25
Db 275 CAGGACAAAGACCGGGAGCTTCCGCGCGGACGTCGCAAAACGCTGATGATCAGAGTT 334
Qy 26 -GluArgGlyGluInlybProProProThrProSerGlyLyseLeu----- 39
Db 335 TGAGGAGATAGAGTCTCTCAGAAATCATCCAGAGCAAACTCTCAGAGTCCCATGGA 394
Qy 40 -----ValSerIleLyseIleGlnMetLeuAspThrGlnG 52
Db 395 GATCGTAAAAAGGCTAAAGACATGACATGCAAGATGACGCTTCGATGATGATCGATGA 454
Qy 52 uAlaPheGluValProGlnArgAlaProGlyLyseValLeuLeuAspAlaValCysAsnH 72
Db 455 CGGCTGATGTATAGAGCGCTCCAGAGCAAGTCTGTTAACAAAGTGCAGACA 514
Qy 72 sLeuAsnLeuValGluGlyAspTYrPheGlyLeuGluPheProAspHislyblyleTh 92
Db 515 TCTGAACTTGCTAGAAAAAGACTCTTCGCTTCACGTATCGAGACCAAGAAATCGAA 574
Qy 92 rValTYrLeuAspLeuLeuysProIleVallyeGlnIleArgArgProlys--HlsVa 111
Db 575 GAACTGGTGGACCTGCTAGAAATTAATAAGACATTCGAAGTGGTGGCAT-- 632
Qy 111 lValVallyePheValVallybPhePheProProAspHisThrGlnLeuGlnGluLe 131
Db 633 ----TTTTCATTAAATGTGAAGTTTAACTTACCAAGACGCTGCTCCAGATTCAGAAATAT 688
Qy 131 uThrArgTYrLeuPheAlaLeuGlnVallyeGlnAspLeuAlaGlnIlyArgLeuThrCy 151
Db 689 CACGAGTACTACTCTGCTTGCAGCTGCGAGATGATATGTGCTGAGCCCTACCTG 748
Qy 151 sAsnAspThrSerAlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPh 171
Db 749 TTCCTTGTAACCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 808
Qy 171 eAsp---GluAlaLeuAspArgGluHisIleAlaValAsnlyblyrIleProGlnGln-- 189
Db 809 TGATCTGATGATGATGTAATAATGATCAATTAAGAGTTCGCTTCCGCTTCCGCTTCC 868
Qy 190 -AspAlaLeuGluuAspLyseIleValGluPheHisIleAsnHisIleGlyGlnThrProAl 209
Db 869 GAAAGAACTGGAGATTAAGTATGATCGAGCTGCACAAAGCCACAGAGAAATGACGCCAGC 928
Qy 209 agIuSerAspPheGlnLeuLeuGluIleAlaArgArgLeuGluMetTYrGlyIleArgyle 229
Db 929 CGAGCAGAGATGCACTTCTGAGGATGCAAAAGCTCCATGTATGGGTAGACTT 988
Qy 229 uHisProAlaLyseAspArgGluGlyThrLybIleAsnLeuAlaValAlaAsnThrGlyI 249
Db 989 ACACCAAGCAAAAGATTCGAAAGCGGTGAGATCATGTATGGGTCTGTGCCACCGGCTT 1048
Qy 249 eLeuValPheGlnGlnlyPheThrLybIleAsnAlaPheAsnThrAlaLyseValArglybLe 269
Db 1049 GTTGATCTATCGTACCGGCTTCGAATTAACAGATTGCTTGCGCCCAAGGCTCTTAATAT 1108
Qy 269 uSerPheLybArgLybArgPheLeuIleLybLeuArgProAspAlaAsnSerAlaTYrGly 289
Db 1109 TTCTTATTAACGGAACAACCTTTTACATTAATATCCGCGCAGAGAGTTGAAACAATTGCA 1168
Qy 289 nAspThrLeuGluPheLeuMetAlaSerArgAspPheCysLyseSerHetrIlyblyleCy 309
Db 1169 AAGCACCATTGGCTTCAAGTTGCCGAACCATAGAGCTGCCAAGGCGCTCGGAAAGATAG 1228
Qy 309 sValGluHisIleAlaPhePheArgLeuPheGluGluProlybProlybProlybProva 329
Db 1229 TGTGGAACATCATCATTTTTCAGACTACG---CTACAGAAAGCACTCCAAAGAAATTT 1285
Qy 329 lLeuPheSerArglybSerSerPheArgPheSerGlyArgThrGlnlybGlnValLeuAs 349
Db 1286 CTG---ACCTGGGCTCAAGTTCCGCTACAGTGGCAGAAACCGACGCGAGAGAG 1342

QY	349	PTyValIysgluGlygluHleValysLysValGlnPheGluArgLysSerValLeuI	369
DB	1343	AGCCAGTCATTTGATGATGAGCCGCCGCCGCTTATTCGAAAGCTGTCACACCAACGATA	14020
QY	369	AserIle---ArgSerLeuAlaSerGlnProThrGluLeuAsnSerGluValLeuGluG	388
DB	1403	CACCATGTCGCCACGCTTGATGAGGCTCAGTGAATGAAAACATGAAATATAC----	1457
QY	388	nSerGlnGlnSerThrSerLeuThrPheGlyGluValaGluSerProGlyGlnSe	408
DB	1457	-----	1457
QY	408	rcYArGArgGlyLysGluProLysValSerAlaGlyLupProGlySer-HisProSerP	428
DB	1458	-----ATGAAGGATCTTATCTGCTGCACAGAGTTGGTATCAGGCCAGTATACG	15040
QY	428	roAlaProArgArgSerProAlaGlyAsnLysGlnAlaAspGlyAlaAlaSerAlaProT	448
DB	1505	CACAAACCAAGGTCATCTCT-CAGACCAACCTGATAC-----ACTGTGACCCCGAGA	15570
QY	448	hrgLugLugLugLugLuvAlValLysAsp-ArgThrGlnGlnSerLysProGlnProPro	467
DB	1558	AAAAGCTGAGAGGAGGACGCGGAGGAAGAGACAGAAAGAAAGAGCCGAGGAAGCCA	16170
QY	468	GlnProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAsnSer	487
DB	1618	CGCCCTGCGAGCCCTCGCGGCGAGGAGAAAGTCAC-----	16530
QY	488	GlnGlyGlyValAlaProAlaAsnValThrLeuSerPro-----AsnLeuSerPro	504
DB	1654	---CAGGGGATGAGCTTCTGATCATGATGCC-CTGTACCTTCATCAGCCCATCTTGACCT	17090
QY	505	-----AspThrLysGlnAlaSerProLeuIle	513
DB	1710	CCCTCTCCACAGAGCTCGTATGAGAGGCAAGAGAAAGAGGCTTCCAC-----	17630
QY	514	SerProLeuLeuAsnAspGlnAlaCysProArgThrAspArgGluAspGlyLysArgLys	533
DB	1764	-----GGTTCTGCTGCTGCTGACGACAGAG-----	17870
QY	534	LysArgPheProThrAspLysAlaLeuYrHeLeuAlaLysGluLysSerThrThGluArg	553
DB	1788	-----CTGTCTCGAGTGTGATCCCAAGGGAG-----	18170
QY	554	ThrTyLeuLysAspLeuGluValIleThrSerThrPheGln-----	567
DB	1818	GCCTACCTCAGGAGCAAGATGTGCATTTGTTACAGACAGCCGCGTGAAGAAAGAAC	18770
QY	568	SerThrValSerLysGluAspAlaMetProGluAlaLeuLysSerLeuIle-----	584
DB	1878	ACACTGTTTCTCTTCTCTGACGCTCCAGAGTCACCCCTGCTCTGATGATGAGAC	19370
QY	585	-----PheProAsnPheGluProLeuHisLysPheHisThrAspPheLys	600
DB	1938	GAGTAACCTCTCTTCTCCCAACCTGCT-----GAAACCAACCTCTGCCC	19820
QY	601	GlnGlnGlnArgLeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIleArgAspLys	620
DB	1983	CAG-----	1985
QY	621	GlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaAlaHis	640
DB	1985	-----	1985
QY	641	LeuTrpLysHisSerGlnAlaLeuGlnAlaLeuGlnAsnGlyIleLysSerSerArgArg	660
DB	1986	AGTTGGCAGCATTTCTTCC-----ATCCGCTGCGCTCTCTC	20240
QY	661	LeuGlnAsnPheCysArgAspPheGluLeuGlnLysValCysTyLeuProLeuAsnThr	680
DB	2025	CTTCCCTGCTCTCTTTCATCTTCT-CTTCCGTGCTGCTGCTCTT-----CTCAGTACC	20800
QY	681	PheLeuLeuArgProLeuHisIleArgLeuMetHisTyTyLysGlnVal-----	695

[illegible]

```
Db 3050 CAAGAGGTGCCAGTGTCCACACAGAGACGAAAACATCACGTACGAGCTTCTCAGGT 3109
QY 969 pAenHiPProleuAlaSerLeuPro-----LeuLeuGlyTyrSerLeuThrIleProSe 987
Db 3110 CGAT--CTGGGGCAGATCTGGAGCCAGGTGTGCTTAATGAGCGCCCAACGATCACGTC 3166
QY 987 rGluSerGluAenIleGlnLysAerTyrValPheLysLeuHisPheLysSerHisValTyr 1007
Db 3167 TGAATCCACGAGTACTACACACACATATCACCAAACTGTGAAGAAGAGGATTC 3226
QY 1007 rTyrPheArgAlaGlu-SerGlu-----TyrThrPheG 1018
Db 3227 GGAGACGAGGATTTGAGAGCGAATAGTCATCAGCGAGATGCCGACATTGACCATGACCA 3286
QY 1018 lAArgTrpMetGluValIle-----ArgSerAlaThrSerSerAlaSerArgProH 1035
Db 3287 GCGCGTGGCTCAGGCAATTAAAGAGGCCAAAGAGCAGACCCCGACATGTCAGTAAACCA 3346
QY 1035 lSValLeuSerHisLysGluSer 1042
Db 3347 A-GTAGTGTGTCATTAAGAGACA 3368
```

Search completed: May 20, 2006, 04:14:36
Job time : 2901 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 20, 2006, 03:27:59 ; Search time 30 Seconds
(without alignments)
3686.546 Million cell updates/sec

Title: US-09-555-342B-2

Perfect score: 5463
Sequence: 1 MGELEQRPTPSRLGAPENS.....SATSSASRPHVLSKESLVY 1045

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 224314 seqs, 35277956 residues

Total number of hits satisfying chosen parameters: 448628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=xlh
-O=/abs8/ABSSWB.epool/US09555342/runat_19052006_121512_13499/app_query.fasta.1
-DB=Published Applications NA New -QFMT=faastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=Blasum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abes02h
-USER=US09555342 @CGN 1 1 13 @runat_19052006_121512_13499 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THRADs=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA_New:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB_seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB_seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB_seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB_seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/PCT_NEW_PUB_seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB_seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB_seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB_seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	713	13.1	2867	US-10-511-937-344	Sequence 344, App
2	396.5	3.3	3953	US-10-473-173-56	Sequence 56, Appl
3	176.5	3.2	6469	US-10-505-928-228	Sequence 228, App
4	173	3.2	2559	US-11-217-529-2747	Sequence 2747, App
5	169	3.1	646	US-10-488-618-2736	Sequence 2736, App
6	155.5	2.8	4119	US-11-217-529-77046	Sequence 77046, A
7	155	2.8	7430	US-10-559-415-1	Sequence 1, Appl
8	152	2.8	7465	US-10-505-928-311	Sequence 311, App
9	151.5	2.8	3054	US-11-217-529-5019	Sequence 5019, App

10	150	2.7	6505	6	US-10-505-928-721	Sequence 721, App
11	140	2.6	7268	6	US-10-559-415-189	Sequence 189, App
12	137.5	2.5	11185	6	US-10-505-928-450	Sequence 450, App
13	137	2.5	11128	7	US-11-118-524-1	Sequence 1, Appl
14	136.5	2.5	5040	7	US-11-217-529-77055	Sequence 77055, A
15	136	2.5	1617	7	US-11-217-529-173406	Sequence 173406, A
16	133.5	2.4	1863	7	US-11-217-529-75500	Sequence 75500, A
17	133.5	2.4	1866	7	US-11-217-529-148	Sequence 148, App
18	133.5	2.4	4080	7	US-11-217-529-1845	Sequence 1845, App
19	133	2.4	3210	7	US-11-217-529-80763	Sequence 80763, A
20	131.5	2.4	4263	7	US-11-327-900-5	Sequence 5, Appl
21	131.5	2.4	5449	6	US-10-505-928-800	Sequence 800, App
22	128	2.3	12515	6	US-10-505-928-358	Sequence 358, App
23	127	2.3	3246	7	US-11-217-529-78477	Sequence 78477, A
24	127	2.3	3318	7	US-11-217-529-4920	Sequence 4920, App
25	127	2.3	3444	7	US-11-217-529-5473	Sequence 5473, App
26	127	2.3	3658	7	US-11-217-529-6026	Sequence 6026, App
27	126.5	2.3	3210	7	US-11-217-529-2695	Sequence 326, App
28	126	2.3	10211	6	US-10-505-928-326	Sequence 326, App
29	125	2.3	3705	6	US-10-505-928-655	Sequence 655, App
30	124.5	2.3	3854	6	US-10-511-937-549	Sequence 549, App
31	122.5	2.2	3261	7	US-11-217-529-1461	Sequence 1461, App
32	122	2.2	2358	7	US-11-217-529-190978	Sequence 190978, A
33	122	2.2	3387	7	US-11-217-529-78908	Sequence 78908, A
34	121.5	2.2	3117	7	US-11-217-529-774	Sequence 774, App
35	121.5	2.2	3729	7	US-11-217-529-79062	Sequence 79062, A
36	120.5	2.2	4053	7	US-11-217-529-53	Sequence 53, Appl
37	120	2.2	4584	7	US-11-217-529-3073	Sequence 3073, App
38	119.5	2.2	9588	6	US-10-505-928-101	Sequence 101, App
39	119	2.2	2667	7	US-11-217-529-2499	Sequence 2499, App
40	119	2.2	3243	7	US-11-217-529-76988	Sequence 76988, A
41	118.5	2.2	3366	7	US-11-217-529-76415	Sequence 76415, A
42	118.5	2.2	3446	6	US-10-505-928-176	Sequence 176, App
43	118.5	2.2	7280	6	US-10-505-928-842	Sequence 842, App
44	118	2.2	2511	7	US-11-217-529-77149	Sequence 77149, App
45	117.5	2.2	3471	7	US-11-217-529-3423	Sequence 3423, App

ALIGNMENTS

RESULT 1
US-10-511-937-344
; Sequence 344, Application US/10511937
; Publication No. US20060088836A1
GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, Macdonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 50661200104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 344
; LENGTH: 2867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-511-937-344
Alignment Scores:

Pred. No.:	1,3e-40	Length:	286
Score:	733.00	Matches:	90
Percent Similarity:	46.1%	Conservative:	190
Best local Similarity:	31.3%	Mismatches:	197
Query Match:	13.1%	Indels:	130
DB:	6	Gaps:	17

```

; NUMBER OF SEQ ID NOS: 498
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 3953
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-173-56

```

Alignment Scores:

Pred. No.:	1 41e-18	Length:	3953
Score:	36.50	Matches:	146
Percent Similarity:	36.4%	Conservative:	98
Best Local Similarity:	23.0%	Mismatches:	245
Query Match:	7.3%	Indels:	147
DB:	6	Gaps:	20

US-09-555-342B-2 (1-1045) x US-10-473-173-56 (1-3953)

QY	29	GlnIysPProPProThProSerIly---	LysLeu-----	Val	40	
Db	197	AAGAAAGAAAGATGCCATCCATTCG	ATTTGGTGAACAGAAACGACCCGGG	CTACAGGTG	256	
QY	41	SerIleIys-----	-----IleGlnMetLeuA	AerPThrGlnIuIaPhe	54	
Db	257	TCCAGCAAGAGATGGCTGGTGGTGG	ATCCCGGATCCAACTGCTTAATAA	AGAGATTGTGGAGTTC	316	
QY	55	GluValProGlnArgAlaProGlyL	ysValLeuLeuA	AlaValCysA	nhiIleuA	74
Db	317	ACCCGTGCTCCGAGAGACATGGCC	ACGAGAAAGCCCTCGAGCCGTG	AGCCGACAGAGCTGGAG	376	
QY	75	LeuValGluGlyAerPThrPheGly	LeuGluPheProAerPheLysVal	IleThrValTrp	94	
Db	377	CTGGGGAGGTCATTA	CTTACCTGACCTCGGTATCA	CAACAGAAATACAGCCCGGTGG	436	
QY	95	LeuAerPLeuLysPProIleValI	ysGlnIle---ArgArgProLys	nhiIleValVal	113	
Db	437	GTAGATTTGGAAAAACCTTGA	AGAGACAGCTGATTAATATGAT	TGGAGACTTACGCTC	496	
QY	114	LysPheValValLysPhePhePro	ProAerPheIsthGlnLeuGlnGlu	LeuThrArg	133	
Db	497	TATTTGGAGTGCGTTTATGTG	CTTCAGTTTCTCAGCTGCAG	ACAGAAATTTCCAGG	556	
QY	134	TyrLeuPheAlaLeuGlnValI	ysGlnAerPLeuAlaGlnI	uArgLeuThrCysA	nAer	153
Db	557	TATCAGTATTAATCTGCAC	AGTAAGAAATATCTTGAAGAA	AGATATCTTGTA	CTTTGACTTA	616
QY	154	ThrSerAlaAlaLeuLeuIle	SerIsthIleValGlnSerGlu	IleGlyAerPheAer	Leu	173
Db	617	GAACAAGACAAATTCAGCT	AGACGCGCTTACTGTCAAGG	GAATTTGGTGA	CTTTGATCAG	676
QY	174	AlaLeuAerPArgGlu	nhiIleuAlaLysA	nLysTrpIlePro-----	Gln	188
Db	677	TATGAATCCGAGCACTTCTT	CAGAAATTTGCTTTGCTG	TCGAGGTGTTACAA	GAT	736
QY	189	GlnAerAlaLeuGluAerP-----	-----LysIleValGlu	PheIsthAerA	nhiIleGly	205
Db	737	GAATAAGATTTGGAGAA	AGAACCAACCAAAAGTGG	CGCTTACTACATCAGAA	TATCAGAGGG	796
QY	206	GlnThrProAlaGluSerAerPhe	GlnLeuLeuGluIleAla	uArgArgLeuGlu	MetTrp	225
Db	797	CTCACAGCTCCGTAGTGA	ATGCTGTATACAGAGAG	GTATGAGAGATGATGG	CTAT	856
QY	226	GlyIleArgLeu	nhiAerProAlaLysAerPArg	GluGlyThrLysIleAer	LeuAlaValAla	245
Db	857	GGAGAAGAGACTTACCTG	CTAAGATGACCAAGAA	AGTGCATATTCATTCAG	AGCGGT	916
QY	246	AerThrGlyIleLeuVal---	-----PheGlnIlePhe	ThrLysIleAerAla	PheAerThrAla	264
Db	917	CTTGAGGTATCTTTGTGA	ACAACAAGATGAGACAG	ATCCGTGTGATTTAG	TGTGGCAT	976
QY	265	LysValAlaArgLysLeuSerPhe	LysArgLysArgPheLeuIle	LysLeuArgProAer	Ala	284
Db	977	GACATTTGCCAACATGTCC	CAACAAGTCTTTTGTGAT	TAGAGACTG-----	GCA	1022

QY	285	AenseralATyrgLnaSPThrLeuGIuPheLeuMetAlaSerATyrgAspPheCysLysSer	304
Db	1028	AATAAA-----GAGGAGACCATTTCAATTCAATTAACCTGAAGACATGAAACACGAAATAATAC	1081
QY	305	PheTrpIlysIleCysValGIuHiHisAlaPhePheArgLeuPheGlu-----	320
Db	1082	ATTGGAGACTCTGTGTGGCCGACACAAAGTTTATACAGACTTAACCAAGTGTAACTGTCAA	1144
QY	321	-----GluProLysPro	324
Db	1142	ACTGAGATGTCAAGTGAACCCATCAAGAGAGAGCTTCTTCAAGATGTCTGTGCT	1201
QY	325	LysProLysProValLeuPheSerArgIlySerSerPheArgPheSerGIlyArg-----	342
Db	1202	AAACCCCGACCTACGTATGATGCTCTCCCAACCGCAGTTGACTATAATGAGCATATATACA	1263
QY	343	-----ThrGIuLysGIuValLeuAspTrpValIlySerGIuGIly-----	355
Db	1262	GAACCATATGCTTCTTCCCAAGATTAACCTTTGTGTCCCAACAGAAACGATTAACCTGT	1322
QY	356	HisIlyLysValGIuPheGIuArg-----LysHisSer	366
Db	1322	CACCTCTCAGACAAAGCTTGATATGAGCCAGATTGACCTCAACGCTGGATCCGTAAATGGC	1381
QY	367	LysIleHisSerIleArgSer-----	373
Db	1382	AGTGTCTACATGTGCACACAGACCAACCTCTTAATAATCTCAGCCCTTAATTCAGAGCC	1441
QY	374	-----LeuAlaSerGIuProThrGIuLeuAsnSerGIuValLeu-----	386
Db	1442	TGCGCGAGTGTCTCCAACTTACGATCAACGGAGTGAACGTATGAGGCTGACTACCTC	1501
QY	386	-----	386
Db	1502	CCGTCCTCATGGCACAGCGCCGTGATACCCCGCTTACCGGCCCAACCCAGACTATATAG	1561
QY	387	-----GluInSerGIuInSer	392
Db	1562	ACTGTGATGAGAGCTCAACAGGGGCGCTGTGCATGCGAAACGGACAGCACTGCTGTG	1621
QY	393	ThrSerLeuThrPheGIly-----	398
Db	1622	CGAAACCTCAACATCGGACGCTCGTACGCTTACAGACAGGCCCGGCGCCTGTGTTACAGC	1681
QY	399	-----GluGIuAlaGIuSerProGIuGIuInSerCysArgArgGIlyLys	413
Db	1682	CAGCCCGAGATCCCGAGACGACGCAAGCTCTCCGCGCAGCGCGCACACTGCCCTGTC	1741
QY	414	GluProLysValSerAlaGIuGIuProGIlySerHisProSerProAlaProArgArgSer	433
Db	1742	AGCTGAGCTACAGCTTCCACAGCCCGTCTCCCTACCTTACCCTGCGGACGGGGGCC	1801
QY	434	ProAlaGIuAsnLysGIuAlaAspGIuAlaAlaSerAlaProThrGIuGIuGIuGIu	453
Db	1802	GTGTGTGGCGCGGTACAGGCTGCCGAGCTGACCAATGGCAGCTGACAGGCCAGACGACTAC	1861
QY	454	ValValIlyAsp-----ArgThrGIuInSerIlySerProGIuProProGIuInProSerThr	471
Db	1862	CCGTCTCCCAACATCATGCGGACGCAAGGTGTACCGGCAACCCCACTTACCTCCGCCCCC	1921
QY	472	GIlySerLeuThrGIlySerProHisLeuSer-----GluLeuSerValAsnSerGIuGIly	490
Db	1922	AGGCCCGCAACAGACGACGACTGTGCTCCGCCACCTTACATCAGAGACGACAC---	1978
QY	491	ValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAspTrpIlyGIuAlaSer	510
Db	1979	-----CCCGACCTCATACCGCGCGGTGACACCACTCGGTGCAAAAGCTTCCAGAGAGAC	2032
QY	511	ProLeu---IleSerProLeuLeuAsnAspGIuAlaCysProArgThr	525
Db	2033	AGCTGCGCGGTGGCAGCTGCTCAGAGAGTGCAGAGCCCTTCACC	2080

RESULT 3

US-10-505-928-228 Application US/10505928
 ; Sequence 228 Publication No. US20060088532A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ludwig Institute for Cancer Research et al.
 ; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
 ; FILE REFERENCE: 28967/39178
 ; CURRENT APPLICATION NUMBER: US/10/505, 928
 ; CURRENT FILING DATE: 2004-08-27
 ; PRIORITY FILING DATE: 2002-03-07
 ; NUMBER OF SEQ ID NOS: 866
 ; SOFTWARE: PatentIn 3.2
 ; SEQ ID NO 228
 ; LENGTH: 6469
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-505-928-228

Alignment Scores:

Prod. No.: 0.00421 Length: 6469
 Score: 176.50 Matches: 160
 Percent Similarity: 36.9% Conservative: 102
 Best Local Similarity: 22.5% Mismatches: 283
 Query Match: 3.2% Indels: 168
 Gaps: 32

US-09-555-342b-2 (1-1045) x US-10-505-928-228 (1-6469)

Qy 311 GlnHisAlaPhePheArgLeuPheGluGluPro-----LysProLysPro-LysP
 |||||
 Db 3436 GAATCGGGTGTGATTTCTTGACCTTGAAAGCGCGGTAGACCAATGCCAGCAATA 3495
 Qy 328 oValLeuPheSerArgGlySerSerPheArgPheSerGlyArgThrGlnLysGlnValle 348
 |||||
 Db 3496 TGTGCTGTTTCGAGCGCAAGCT-----AAGCAGGGCGCT 3528
 Qy 348 uAPRYValLysGluGlyGlyHisLysValGlnPheGluArgLysHisSerLys11 368
 |||||
 Db 3529 TGACTGATCCAGAAACAGGT-----GAATTT-----TACCTCTTAAC 3567
 Qy 368 eHisSerLeuArgSerLeuAlaSerGlnProThrGlnLeuAlaSerGlnValleuGlu1 388
 |||||
 Db 3568 ACATACCTCCACGAGAGACACAGAGAGACTCAGGA-ACGCTGAAAGATATGGGG 3626
 Qy 388 nSerGlnGlnSerTherSerLeuThrPheGlyGlyGly-----AlaGluSerProGly1 406
 |||||
 Db 3627 AATTCAGGGTGCCTGCCAGCAACAAAGGAGAGGTAAGCTTGATTCAGCTGGCCG 3686
 Qy 406 yGlnSerCysArgArgGlyLysGluProLysValSerAlaGly-----GluProGlyse 424
 |||||
 Db 3687 ATAGCTTTGT-----GAAHAAAGCCACATTCATGCCACGAGATAGAAATGGGTGA 3740
 Qy 424 rHisProSerProAlaProArgArg--SerProAlaGlyAlaLysnLysGlnAlaaspGlyAla 443
 |||||
 Db 3741 CCACGGGCAAGACACTACAGAGATTTCTCCCTGAGAGATGGAAATACCAATACAC 3800
 Qy 444 AlaserLalProThrGluGluGluGluValValLysAspArgThrGlnGlnSerLys 463
 |||||
 Db 3801 TGGAGAAAGCCCTTAGAGTCAACACAGAGATAT--AAGGAC----- 3841
 Qy 464 ProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeu 483
 |||||
 Db 3842 -----CTGAGACTG 3850
 Qy 484 SerValaenSerGlnGlyValAlaProAlaAsnValThrLeuSerProAsnLeuSer 503
 |||||
 Db 3851 GAT-----ATTATCCAGCAAGC----- 3668
 Qy 504 ProAspThrLysGlnAlaSerProLeuLysSerProLeuLeuAsnArgGlnAlaCysPro 523
 |||||
 Db 3869 -----CTTTCGATCGGAGGTCAAG 3889

Qy 524 ArgThrAspArgLysAspGlu-----GlyArgArgLysArgPheProThrAspLysAlaLysr 542
 |||||
 Db 3890 CTCGGGAGCCCAACACAGATCATGAAGAACCGGAAGTCAACCCCGGAAGAAAGA 3949
 Qy 543 PheLeuAlaLysGluValSerThrThrGluArgThr-TyLeuLysAspLeu---GluVal 561
 |||||
 Db 3950 TTTATATATGCTGATCACTCCAGACAGAGAGGCTTATGTAAGGATTTGATGAGTGC 4009
 Qy 562 IleThrSerThrPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeu--- 580
 |||||
 Db 4010 TTAGAGACCTACCTGGGAAATGACACAGTGTGTGAGAGAGATCCCTCGGATCTC 4069
 Qy 581 -----LysSerLeuLysPheProAsnPheGluProLysLysLysPheHisThrAsn--- 597
 |||||
 Db 4070 AATTAAGAGCATATCATCTTTGGCAACATCCAAAGAGATCTACGATTTCCATACACATC 4129
 Qy 598 PheLeuLysGluLysGluGlnArgLeuAlaLeuTyrglu-----GlyArgSerAsnAla 615
 |||||
 Db 4130 TTCTCAAGAGCTGAGAGAAATGACGAACTCCGAGGATGTGGACACTCTTTGTT 4189
 Qy 616 GlnIleArgAspTyrglnArgLysGlyAspValMetLeuLysAsnLysGlnLysMetLys 635
 |||||
 Db 4190 ACCTGGGACAGCAAAATTTACAGATGATGTCACTACTGTAAACAAAGCTGATCCAC 4249
 Qy 636 HisLeu-----AlaAlaHisLeuTyrglyHisSerGluAlaLeuGluAlaLeu 651
 |||||
 Db 4250 CAGCTTATCTGAGGATGCGGACCTCTTTGATGAGATACACAGCGGACGTGCTG 4309
 Qy 652 GluAsnGlyLysSerSerArgLysGluGluAsnPheCysArgAspPheGluLeuGln 671
 |||||
 Db 4310 GCCAACTCCATCTCTTC----- 4327
 Qy 672 LysValCysTyrglyLeuProLeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHis 691
 |||||
 Db 4328 -----TACCTATTAAGCTCTCCAAAGATCAACAA 4360
 Qy 692 TyrglyGlnValLeuGluArgLeuCysLysHisProProSerHisValaAspPheArg 711
 |||||
 Db 4361 TATCACTGCTCTGAAAGAACTTTAATCTGCTGAGAGAGAGGAGGAGCTCAAG 4420
 Qy 712 AspCysArgAlaAlaLeuAlaGluLysThrGluMetValAlaGln---LeuHisGlyThr 730
 |||||
 Db 4421 GATGGCTGAGGATGATGCTCAGTGTCCCAAAAGGACCAATGATGCTGATGCTGACG 4480
 Qy 731 MetLysMetGlnAsnPheGlnLysLeuHisGluLeuLysAspLeuLysGlyLe 750
 |||||
 Db 4481 ATG-----CTGAAAGGTTCCAGAGAACTGAGATGCGAGGGAGTTGATTTCCAG 4534
 Qy 751 AspAsnLeuValVal---ProGlyArgGluPheLysArgLeuGlySerLeuSerLysLeu 769
 |||||
 Db 4535 GATGCTTTCAAGTGTGGAGCCGAGAGTGTGATCCGGAAGGG----- 4579
 Qy 770 SerGlyLysGlyLeuGlnGlnArgMetPhePheLeuPheAsnAspValLeuLeuTyrgThr 789
 |||||
 Db 4580 -----CGGAGCGGCACTTGTCTCTTGAGATCTCTTGATTTTAAAG 4624
 Qy 790 SerArgGlyLeuThrAlaSerAsnGlnPheLysValHisGlyGlnLeuProLeuTyrgGly 809
 |||||
 Db 4625 AAGAGATCAAGATCTTCAAGACACAGAAATAGTTTACAAAGAAAGCTTACGACC 4684
 Qy 810 MetThrLysGluGluSerGluAspGluTyrglyValProHisCysLeuThrLeuArgGly 829
 |||||
 Db 4685 TCAGACTGGGTGTACCGAGACACGAGGAGGATCCCTGCAAAATTCGCTGTGTCT 4744
 Qy 830 GlnArg-----GlnSerLysLeuAlaAlaAspSerArgSerGluMet 844
 |||||
 Db 4745 GGGCGACACCCATCTCAGACATTAAGACAGCTGAAAGCTTCAACATTAAGAACCAAG 4804
 Qy 845 GluLysTyrgValGluAspLysGlnMetAlaIleAsp-----LeuAlaGlyLysSer 861
 |||||
 Db 4805 CAGAGTGTATCAAGAACATTGAGAAAGTGAATTCAGAAAGATCATTCACCTGAAAGGA 4864

```

Oy      862  SerSerProAlaProGluPheLeuAlaSerSerProPro-----AspAsnLysSer  878
Db      4865  GCTTTAAAGAGACCCTTCAGCTCCCCCAAAACACAGCCAAACAGAGAAACAATGTGAAG  4924
Oy      879  ProAspGluAlaThrAlaAlaAspGlnLysSerGlu-----AspAsnLeu  893
Db      4975  AGGAGTGAAGTGGAGAGATTTCACACAGCCAGGGGAGATGGAGACGCCAACACACACCTTC  4988
Oy      894  Ser---AlaSerArgThrSerLeuGlnArgGlnAlaProHisArgGlyAsnThrMetVal  912
Db      4985  TCCATTGGCTCTTAGGACCTCTCAGAAACACAGTGCACAGTGCACAGAGATGGCAACTTGTT  5044
Oy      913  HisValCysThrPheIleArgAsnThrSerValSerMetValAspPheSerIleAlaValGlu  932
Db      5045  ---CCTCGGTGGCACCT--GGAGACCTGAGATCTTTCTCCATTACGTTTAGCG-----  5094
Oy      933  AsnGlnLeuSerGlyAsnLeuLeuArgLysPheLysAsnSerAsnGlyTTPGlnLysLeu  952
Db      5095  ---GCCATCTCGGGAGCTTGTCCTCTCAGACTACCGGGTTTAGCGGTGGC-----  5139
Oy      953  TrpValValPheThrAsnPheCysLeuPhePheTyrLysSerHisGlnAsp---AsnHis  971
Db      5140  ---ACCGTTGGAGCC-----TCCACCAAGACATCCCATATCAC  5172
Oy      972  ProLeuAlaSerLeuProLeuLeu  979
Db      5173  CCA-----CCTCTCTTG  5184

RESULT 4
US-11-217-529-2747
; Sequence 2747, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKMO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 6-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 2747
; LENGTH: 2559
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-2747

Alignment Scores:
Pred. No.:      0.00215      Length:      2559
Score:          173.00      Matches:      114
Percent Similarity: 38.04      Conservative: 83
Best Local Similarity: 22.04      Mismatches: 187
Query Match:    3.24      Indels:    137
DB:            7      Gaps:    23

US-09-555-342B-2 (1-1045) x US-11-217-529-2747 (1-2559)
Oy      482  GluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerProAs  501
Db      681  GACCTCTCAATTGGTCAAAAGTGGAGGTAGTGAACCGCTACGATTCACAGTCCAA  740
Oy      501  nLeuSerProAspThrLysGlnAlaSerProLeuLysSerProLeuLeuAsnAspGlnAl  521
Db      741  TATCTTCCCTCCAGAGACAAACACAGCAAACTTG-----  777
Oy      521  aCysProArgThrAspAspGluAspGlnGlyArgArgLysArgPheProThrAspLys--  540

```

Db	778	-----ACTGCAGAAAATCAACAGCATCTTCAACAGCATGCTCTTCAAAAAACA	827
Qy	541	----AlaTyrPhe---IleAlaIysGluValSerThrThrgluArgThrTyrLeuIysAs	558
Db	828	TATTGAGTACGTTAAGATTATTCACAAAGAAATTGGTGCACACAGAAAGAAAGTACGTTCA	887
Qy	558	PLEUGLIVallIleThrSerTrpPheGlnSerThrValSerIySgluAspAlaMetProG	578
Db	888	TTTAAAGATTTTGGACAAAGTATAGACCAACAGCTATTAGACAGTGCACCTGATCAGCTCGA	947
Qy	578	uAlaLeuIysSerLeuIlePheProAnPheGluProLeuIleIySyrPheIleThrAspH	598
Db	948	AGAGCTT---TATATGTTGTTCCCACTTGAAGCAGCCGATAGATTTTCAAGAAAGATT	1004
Qy	598	eLeuIySgluIleGluGlnArgLeuAlaLeuTrpGluIyArgSerAnIaGlnIleAr	618
Db	1005	TTTAGTATCCCTGGAG-----ATCAATGCTCTTAGTGA	103
Qy	618	g---AspTyrGlnIyArgIleGlyAspValMetLeuIySasnIleGlnIyMetIySAla	637
Db	1038	ACCTTCTAAGCAGAGAACCGGTGCTTTTCATGCAT-----	107
Qy	637	uAlaAlaIleIleuTrpIyS-----HisSerGluAlaLeuG	649
Db	1075	---TCCAAAATTTTTCAGACTTATAGCCTTGCTGCATTTGCTCAAAATGGCGGCATGCA	1133
Qy	649	uAlaLeuGluAnGlyIleIySerserArgArgLeuGluAnPheCySArgAspPheG	669
Db	1134	GTTTCTTACTTCCACTTA-----AACAAATGAGATCTGATGA	1177
Qy	669	uLeuGlnIySValCySTr-----LeuProLeuAnThrPheLeuLeuArgProle	686
Db	1173	ATCGCAGAACTTCAATATCAACAATTAACCTAGAAATTGACAGTCTTCTCTAATGACCGGT	123
Qy	686	uHisArgLeuMetIleTyrIySglnValLeuGluArgLeuCySlySyrHisIleProPse	706
Db	1233	GCAACGGCTCGCAGGATCTCTTAATGATTAAGAAATTTGGTGGCTTACTTAGATGA	1295
Qy	706	rHisAlaAspPheArgAspCySArgAlaAlaLeu-----	717
Db	1293	CAAC---AATTCAMAGAACTCGAAAGCTGCTTGATATCTTCACAAATATCGCAAGAG	1344
Qy	718	-----AlaGlnIleThrGluMetValAlaGlnLeuHisGlyTh	730
Db	1350	TATCAACGAGAACCAAGAAAGAAAGACTGAAATCACCAAGTGTATAAAAAATCTTATGTGAG	1405
Qy	730	rMetIleIySMetGluAnPheGln-----IySLeuHisGluLeu-----	744
Db	1410	AGTGTAATTTGGAAGGGCTTAAAGATCTCTAAGTTGGTGAATTAATTTAGTACTTCGATTA	1465
Qy	744	uIyAspAspLeuIleGlyIleAspAnLeuValProGlyArgGluPheIleArgLeuG	764
Db	1470	AGTGTCAATTTCCACAAACAACAAATTTCTACGAACTCGAAAGAGATTC-----	1518
Qy	764	ySerIeuserIySLeuSerGlyIySgIyLeuGlnIyGlnArgMetPhePheLeuPheAsnAs	784
Db	1519	-----GAAAGTTATCTTTTGAAGAAAATTTATCATATGTTTTCGGA	1555
Qy	784	rValIleuLeuTyrThrSerArgIyLeuThrAlaSerAnGln-----PheIySValHis	802
Db	1560	AGTATAGTACCTAAGAAATCCACATC-ATCCATTAATGCTCAACAAAAAATCTTCCGCTCAG	1618
Qy	802	sgIyGlnIeupProLeuTyrGlyMetThrIleGluIuseuIyAnGluuTrpGlyValPr	822
Db	1619	CAT-CAATGTCA-----ACTCGGAGCATAGCAACAAT	1653
Qy	822	ohiCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaIaIaSerArgSe	842
Db	1654	AACAGTCTCACCATAATATATCATAGAGAGACGCAATAGTACT-AGTATATATATAT	1712
Qy	842	rGluMetGlnIySTrpValGluAspIleGlnMetAlaIleAspLeuAlaGluIySers	862
Db	1713	CCATTTA-----TCCTC	1724

QY 862 rSerProAlaProGluPheLeuAlaSerSerProAlaPheLeuSerProAlaGluAl 882
DB 1725 TTGCTGACGACGGCGCATTAATAATTCAT-----AT 1757
QY 882 aThAlaAlaPheGlnGlnSerGlnAspAspLeuSerAlaSerArgThrSerLeuGluAr 902
DB 1758 CAACGGCGGTATCAATTCACAAATTGCTTATTCATTCATTAATATATACAC 1817
QY 902 GGLAlaProHis-----ArgGlyAsnThrMetVal-----HisValCy 915
DB 1818 CAATGAACCTTAAGTCTGATGAGAGGTAGATTAATGATTAATGAATCAAAATTAAT 1877
QY 915 sTrpHisArgSerThrSerValSerMetValAspSerIleAlaValGluAsnGlnLe 935
DB 1878 ACCCAAAACACAGATCATTAACATTAAGTCTGGAATCCATAAAGACA----- 1929
QY 935 uSerGlyAsnLeuLeuArgGlyPheLeuAsnSer-----AsnGlyTrp 949
DB 1930 ----GGAATTTCTTTAAAGTTCAAAAATGAAAGAAACGAGATTAATTGG 1977

RESULT 5
US-10-488-619-2736
; Sequence 2736, Application US/10488619
; Publication No. US20060099578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P.C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US/10/488, 619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2736
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-2736

Alignment Scores:
Pred. No.: 0.000659 Length: 646
Score: 169.00 Matches: 47
Percent Similarity: 50.9% Conservative: 38
Best Local Similarity: 28.1% Mismatches: 72
Query Match: 3.1% Indels: 10
DB: 6 Gaps: 4

US-09-555-342B-2 (1-1045) x US-10-488-619-2736 (1-646)
QY 660 ArgPheGluAsnPheCybArgAspPheGlu-----LeuGlnLysValCysArg 675
DB 32 AAGGTTCAGCAAGTACGTACTTCTTGAAGGCTCCGCGCTGCTGCAGAAAGGATATTGAC 91
QY 676 LeuProLeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisArgLysGlnVal 695
DB 92 ATCTCGGTGATGCTTCTTCTGCTGACCCCGGTGCAAGAAATCTGCAAGTATCCCTCAG 151
QY 696 LeuGlnArgLeuCybLysHisArgProProSerHisAlaAspPheArgAspArgAla 715
DB 152 CTGGCGGAGCTGCACAGTACACACACCCTGACAGGAGCTTTAAGAAATGTTGAAGCT 211
QY 716 AlaLeuAlaGluIleThrGluMetValAlaGlnLeuHisArgLysThrMetIleLysMetGlu 735
DB 212 GCTTTCACCGCATGAAGATGTGCCCCAGCTCATCAACGAAACGAAACGAACTTGA 271
QY 736 AsnPheGlnLysLeuHisArgLysLys-----LysAspLeuIleGlyLysAspAsn 752
DB 272 AACATCGACAAGATGCTGCTGACAGAGCTCATAGAGAGCTGGAGGGG---GAGAC 328
QY 753 LeuValValProGlyArgGlnPheIleArgLeuGlnLysLeuSerLysLeuSerGlyLys 772
DB 329 CTCTGTGTGAGAGCTCGAACTCATCTCGGGGAGGCTGACCCGCTGACACAGCCT 388

QY 773 GlyLeu-----GlnGlnArgMetPhePheLeuPheAsnAspValLeuLeuTyThrSer 790
DB 389 CAACCGAGAGTACGACAGAGATGTTTTCTCTTTCACCGCCAGCTCATCTACTGAAA 448
QY 791 ArgGlyLeuThrAlaSerAsnGlnPheLysValHisGlyGlnLeuProLeuTyGlyMet 810
DB 449 AAGGATCTGCTGCGCGGAGGTGATGATACAAAGGCGCGCTGACATGATGACCTG 508
QY 811 ThrIleGlnLysSerGluAsp 817
DB 509 GAGGTGTGATGATGTGAAGAT 529

RESULT 6
US-11-217-529-77046
; Sequence 77046, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIMIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217, 529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77046
; LENGTH: 4119
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-77046

Alignment Scores:
Pred. No.: 0.0653 Length: 4119
Score: 155.50 Matches: 190
Percent Similarity: 37.8% Conservative: 140
Best Local Similarity: 21.8% Mismatches: 321
Query Match: 2.8% Indels: 223
DB: 7 Gaps: 49

US-09-555-342B-2 (1-1045) x US-11-217-529-77046 (1-4119)
QY 22 IleSerThrLeuGlnArgGlnLysProProThrProSerGlyLysLeuValSer 41
DB 1111 ATCAATTTTACAGATGCGCTTAAGATGATCTTAATAGTGTGTCGATCATGAGACAA 1170
QY 42 IleLys-----IleGlnMetLeuAsp 48
DB 1171 ATAAACAAATTCAAATCCAAAGATCTGCATATTGACTCGATGATAAAACATACCAAC 1230
QY 49 AspThrGlnLysAlaPheGlnValProGlnArgAlaProGlyLysValLeuLeuAspAla 68
DB 1231 GATACGAATTCGCTTAC-----CCAAACAAAGATATGCCAAAGATAGAAAGGAGCCT 1284
QY 69 ValCys---AsnHisLeuAsnLeuValGlnGlyAspTyPheGlyLeuGlnPheProAsp 87
DB 1285 CTATGCTTGAATAATGCTGTTATTGAGGCAAAAGATAGTCCCGTAGAGGCCAATA 1344
QY 88 HisLysLysIle---ThrValTrp-----LeuAspLeuLeuLysProIleValLysGln 104
DB 1345 AACGAATTAATTAATTCGCTGTGAAATCTTAGACTTCAAAACCATACTCGAATCT 1404
QY 105 IleArgArgProLys-----HisValValLysPheValVal 117
DB 1405 ATTAACTATTAAATTTGATTAATCTTTGCTGTTTATATCTCATAGATTCCTTCAAGTG 1464
QY 118 LysPhePheProPro---AspHisThrGlnLeuGlnGlnLysLeuThrArgTyLeuPhe 136

[illegible]

Db	2365	GAAGCTAAACCGGATGCGTGTCTCTGATTTGCAGACACTGCTCG--	CCAACCCCTTCCA	2421
Qy	4453	GIUVALValyAspArgThrGlnGlnSerLysProGlnProGlnProSerThrGly	4472	
Db	2422	GATCTATTAAACCTAAACCTTGGCCGCTCCACACACCG-----	2466	
Qy	473	SerLeuThrGlySerProHisLeuSerGlnLeuSerValAanSerGlnGlyValAla	492	
Db	2467	-----CCTTTACCCGAGCTCTTTCGATGAACAAAGAC-----	2499	
Qy	493	ProAlaAnValThrLeuSerProAnLeuSerProAspThrLysGlnAlaSerProLeu	512	
Db	2500	CCCTCCAAATCAGATTATGTT-----ACTCTCCAGCTCCACTTTTA	2541	
Qy	513	IleSerProLeuLeuAsnAspGlnAlaCys-----ProArgThrAspAspGlnAsp	529	
Db	2542	CCAAATGGCTCTTCTTCTCTCCCTCCAGTCGTATCAATCCACAGCAACCGAT-----	2595	
Qy	530	GluGlyArgArgLysArgPheProThrAspLysAlaTryPheIleAlaLysGluValSer	549	
Db	2596	-----TTAAACCACTCCAACTGAAGCGA-----TTGAACCAATCCAC	2637	
Qy	550	ThrThrGluArgThrTryLeuLysAspLeuGluValIleThrSerTrpPheGlnSerThr	569	
Db	2638	TGGATGAAGTGGAGGATATCAAGAC-----ACACTTTGG--GAAGACCG	2682	
Qy	570	ValSerLysGluAspAlaMetProGluAlaLeuLysSerLeuIlePheProAnPheGlu	589	
Db	2693	TTTCAACCCCAAGAACATCAAAAGAAATTACAACTGATGTAATTTCTCAGATTGAA	2742	
Qy	590	ProLeuHisLysPheHis-----ThrAnPheLeu-----LysGlnIleGluGlnArg	605	
Db	2743	GATATTTTAAAGATGAAGAAAGTCCGACTAAGTGTGAATTAAGAAAGAACGCAAGATCTCA	2802	
Qy	606	LeuAlaLeuTrp-----GluGlyArgSerAnAlaGlnIle-ArgAspTryGlnAlaGyl	623	
Db	2803	ATTGGCTGTCTCTCAACACATGGGACAGTCTTCATGACAAACGAAAGAAATTTCTATCTTA	2862	
Qy	623	eglYAspValMetLeuLysAsnIleGlnGlyMetLysIleuAlaHisLeuTrpLys	643	
Db	2863	TCAGAGATTTGGCTCGACGCAATTTGGTATTAATTTA-CACATGTTTCCCAATTATCTGA	2921	
Qy	643	HisSerGluAlaLeuGluAlaLeu-----GluAnGlyIle--LysSerSerArgAr	660	
Db	2922	TATGGAATTTGTATGAAGAATTAATGAACGTGATGATACAGCACTTTTCAAAATGTTAAAT	2981	
Qy	660	GLeuGluAnPheCysArgAsp-----PheGlnLeu-GlnLysValCysTryLeuProL	678	
Db	2982	ACTGAAATTTTGTGTAAAGAAAGATTAGTAATATACCAAAAGTATGCTTAATAGTA	3041	
Qy	678	euaenThrPheLeuLeuArgProLeuHisArgLeu-----MetHisT	692	
Db	3042	CGAGC-----CATATTCACAGGATGAAGATGTTAAACGACTAAAGTCACTT	3086	
Qy	692	YrLysGlnValLeuGlu-----	697	
Db	3087	ACAAAGACCTACACGAATATTTTGGAGCTGTGATCAATTTAAAGATTTTATTTGGAATGT	3146	
Qy	698	-----ArgLeuCysLysHisHisProProSerHisAlaAspPheArgAspCysArgA	715	
Db	3147	AAGATCAAAAGAGCTGCTGCATATTATACATACAGAGAGA-GATTATTAGCATTTGATTT	3205	
Qy	715	IaAlaLeuAlaGlnIleThrGlnMetValAlaGlnLysHisGlyThrMetLysMetG	735	
Db	3206	TCAAAGTTCAAAAAATGATGATCAATTCACACTTAAT-----CGTTTAC	3253	
Qy	735	IuaenPheGlnLysLeuHisGlnLeuLysLysAspLeuIleGly-----IleAspAnL	753	
Db	3254	CTAAATTTTAAAGTTTGAATTTATTTATTTACAGAA--ATAGCAATCATATGAATAA	3310	
Qy	753	euaValValProGlyArgGluPheIleArgLeuGlySerLeuSerLysLeuSer-----G	771	
Db	3311	GAATTTTAAAGGCT-----ATCAAAATTTGAAGTCAATTAACCTTGGTGTCA	3361	

```

RESULT 7
US-10-559-415-1
; Sequence 1, Application US/10559415
; Publication No. US20060100132A1
; GENERAL INFORMATION:
; APPLICANT: Astrazeneca AB et al
; TITLE OF INVENTION: Diagnostic Method
; FILE REFERENCE: 101073-1P WO
; CURRENT APPLICATION NUMBER: US/10/559,415
; CURRENT FILING DATE: 2005-12-06
; PRIOR APPLICATION NUMBER: 031081.2
; PRIOR FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 7430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-559-415-1

```

Alignment Scores:

Pred. No.:	0.154	Length:	7430
Score:	155.00	Matches:	128
Percent Similarity:	35.1%	Conservative:	82
Best Local Similarity:	21.4%	Mismatches:	221
Query Match:	2.8%	Indels:	167
DB:	6	Gaps:	22

US-09-555-342B-2 (1-1045) X US-10-559-415-1 (1-7430)

QY 375 A l a s e r g i n P r o t h r g i u l e u -----A a n s e r g i v a l l e u g i n s e r g i n 391
||| : : : : :
Db 7 G C G G C G G G C C C C C A C C A T G A G G A G C C C C A G C G C G G S A G C T G C C C A G T C C A G C A G 66
392 S e r T h r S e r t u e u r h r h e g i y g i u -----G l y 400
||| : : : : :
QY 67 A G C C T G G C C C A G G C C A T A C G A G A G T G A A G C C G T G C T G G G C T G C T G A G G C C G G A 126
401 A l a g i u s e r - P r o -----G l y e l y i n s e r C y a r g h a r g e l y i u g i u p r o l y s v a l s e 418
127 G C G C T A G T C C C G G C A G C G G C G G C A G T G A C G A G A G A G C G G A G C G C C A A G C G A G 186
QY 418 F A I A G l y i u P r o g l y S e r H i s P r o - S e r P r o A l A P r o A r g A r s e r P r o L a c l y s n L 438
187 C T G C T G C T A A G C T G C T T T T G G C A A G A G C G G A A C C T T C C A G A C T T G C G G C G C G 246
QY 438 Y e g i n a l a e r g i y a l a l a s e r a l a P r o T h r g i u g i u g i u g i u - V a l l y a s p 457
247 C T G G A A A G A C G A G C C T A C C T G C G C C A T T C T A C T G A A C G G G T G T C ----- 300
QY 458 A r g T h r i n g i n s e r l y s e r P r o g i n P r o P r o g i n P r o s e r T h r i g l y s e r l e u T h g l y s e r 477
301 -----G G G C C G C G C A G C C C G C G A A G G C G G T T T A C C T A C 339
QY 478 P r o H i l e u s e r g i u l e u s e r V a l a n s e r g i n g l y i V a l l a P r o A l A s n V a l T h r 497
340 A G C G C T T G T C C A C C A T G C C C T C A G A C T A G A A A C A G A C A G C T C C T C A G A G T G --- 386
QY 498 L e u s e r P r o A n l e u s e r P r o A s p T h r l y s g i n a l a s e r P r o l e u l l e s e r P r o l e u l e u 517
397 -----G G C A C T A C C G G A A G G C G C G T C C C A C C C C T C T C 435
QY 518 A s n A s p g i n l a c y b P r o A r g -----T h r 525
436 A C T G A C C A G A C T G A A T A G A A A G T G A A A C T T C C A T T A C A G C T G C G G C T A T A C C 495
QY 526 A s p A s p i u s e r g i u l y a r g A r g l y s ---A r g P h e P r o T h r A s p l y s a l a t y P h e I l e 544

Dd	496	CGGAGAGAAACGAGCTCCGCAAGCCGTGGACCTTTCTCAAGCATGGACGAGCTTTGAC	555
Oy	545	AlaLeuGluValSerThrThrGluArgThrTyr-----	555
Dd	556	AAAGAGCCCTAACCAAGAGCTGAATCTCGATCATGAGAGCTGAAGATCAATGACGTCGCGA	615
Oy	556	---LeuLysArgLeuGluValIleThrSerThrPheGlnSerThr-----	569
Dd	616	GCCATGTCGCGACCTCGCAGAGCTGTGCAGAACCCAGACACCAACGCTTGAAGAGTGTAG	675
Oy	570	---ValSerLysGluAspArgIleMetProGluAlaLeuLysSerLeuIlePhe-----	585
Dd	676	GAGGTGGCCAGAGAGATGACTTTCTACCAACACTCCACAGCCGGAGCTTCGATGACACAG	735
Oy	586	-----ProenPheGluProLeuLysIleValPheIleValThrAspPheLeuLysGlu	601
Dd	736	ACCTGGCTGAAGAGATGACGTTGACATGCTGAAGCGGAGAAATGGAGACGCTGTGGGAG	795
Oy	602	IleGluGlnArgLeuAlaLeuThrGluGlyArgSerAsnAlaGlnIleArgAspTyrGln	621
Dd	796	CGAAACCTGTGCAGCAGTATGGAGAGATGACAGCGGCTCCACAGAGAGACCAAGAG	855
Oy	622	ArgIleGluAspValMetLysLysAsnIleGlnGlyMetLysLysIleAlaIleAsnIleLeu	641
Dd	856	GAGATCCGATGACCTCCCTGCCACAGACAGACAGAGGTGTGAAGCCAAACGGCTG-----	909
Oy	642	TrypHisSerSerGluAlaLeuGluAlaLeuGluLysIleLeuLysSerSerArgArgLeu	661
Dd	910	-----TCCAGATTTCTCAACAACTGTATGAC-----ACGGCATGACAACTTG	954
Oy	662	GluAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuProLeuAsnThrPhe	681
Dd	955	GAGGTGGTCAAGAAAGCATGACAGCCCTTCGGAAGAGGTACAGTGAAGAAAGTGGCATC	1014
Oy	682	LeuLeuArgProLeuHisArgLeuMetHisTyrTyrLeuIleValLeuGluArgLeuCysLys	701
Dd	1015	CACAAATCGAAGCTGAGCCGCTGGAGACAGCTGGAGGAGAAACCAAGCGTTGTAGG	1074
Oy	702	HisIleProProSerHisValAspPheArgAspCysArgAlaAlaLeuAlaGluIleThr	721
Dd	1075	CAG-----ACA	1080
Oy	722	GluMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsn-----	736
Dd	1081	GAGATGCTGACCCGACAGAGGACACAGCGCATCCAGCTGACAGCACCAGTCCGCTCTCC	1140
Oy	737	-----PheGlnLysLeu---HisGlnLeuLysLysAspLeuIleGlyIleAspAsn	752
Dd	1141	CTGAGAGAGTTTGAAGGAGATCCACCAAGACTGAACAG-----	1179
Oy	753	LeuValValProGlyArgGluPheIleArgLeuGlySerLeuSerLysLeuSerGlyLys	772
Dd	1180	-----GCCAGGGCCGACACAGAG-----	1197
Oy	773	GlyLeuGlnGlnArgMetPhePheLeuPheAsnAspValLeu---LeuTyrThrSerArg	791
Dd	1198	GACCTGACGTGGAGATGAGAGCTGCTCAGTCAAGAGCTGACCAAGACCAACACAGCAG	1257
Oy	792	GlyLeuThrAlaSerAsnGlnPheLysValHisGlyGlnLeuProLeuTyrGlyMetThr	811
Dd	1258	GTGAAGACAGAG-----	1269
Oy	812	IleGluGlnSerGluAspArgIleThrGlyValProHisCysLeuThrLeuArgGlyGlnArg	831
Dd	1270	---AAGAGCTCGAG-----AAATGACAGGAGAGACGG	1299
Oy	832	GlnSerIleIle-----ValAlaLaserSerArgSerGlnMetGluLysTyr	847
Dd	1300	GACGCTGTGAACAGCGATCAAGCTCATATGATGAGGTGACCAAGTCAATCTCGAG	1359
Oy	848	ValGluAspIleGlnMetAlaIleAspLeuAlaGluLysSerSerSerProAlaProGlu	867
Dd	1360	CTGGAACAAGCTGCAGACCGAAGTGGAGACTGGCCGAGTCCAG-----	1401


```

Oy 868 PheleuAlaSerSerProProAbaPnlySerSerProAbaPnlyAlaThrAlaAlaSerGln 887
Db 1402 ---CTCAAGACACACATCTTGAGAAAGGCGGCCATGAGAGATGAGGCGCTGGCG 1456
Oy 888 GluSerGluAap-----AapLeuSerAlaSerArgTrpSerLeuGlu 901
Db 1459 CAGATCAAGACACGCTGACATGATGATCTGGAGAGACCAACAGAG 1506

RESULT 8
US-10-505-928-311
Sequence 311, Application US/10505928
Publication No. US20060088532A1
GENERAL INFORMATION:
APPLICANT: Ludwig Institute for Cancer Research et al.
TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
FILE REFERENCE: 28967/39178
CURRENT APPLICATION NUMBER: US/10/505,928
CURRENT FILING DATE: 2004-08-27
PRIOR APPLICATION NUMBER: US 60/363,019
PRIOR FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 866
SOFTWARE: PatentIn 3.2
SEQ ID NO 311
LENGTH: 7465
TYPE: DNA
ORGANISM: Homo sapiens
US-10-505-928-311

Alignment Scores:
Pred. No.: 0.25 Length: 7465
Score: 152.00 Matches: 201
Percent Similarity: 31.0% Conservative: 118
Best Local Similarity: 19.5% Mismatches: 339
Query Match: 2.8% Indels: 373
DB: 6 Gaps: 47

US-09-555-342B-2 (1-1045) x US-10-505-928-311 (1-7465)
Oy 152 AenAapThrSerAlaAlaLeuLeuLeuSerHisAlaValGlnSerGluTleGluAapPhe 171
Db 1644 AATGAGCACTGCAAGCAAGTCTTGTGCGGCACTGTTCAACAGTGCAGCAAGGAGAAAT 1703
Oy 172 Aap---GluAlaLeuAapArgGlnHisLeuAlaValAaValyetyTleProGlnAap 190
Db 1704 GACCTGGAGAGATGATGACTGGCTGCAATC-----GAGTTCACTGCAACAAGAT 1755
Oy 191 AlalaLeuGluAapLeuTleValGlnPheHisValSerHisAlaGlnThrProAlaGlu 210
Db 1755 GCCCTG---GACATGATTTGCCAACAAGCCCAATGAAATCATCTCCCTCATGATGAGAG 1812
Oy 211 SerAapPhe-----GlnLeuLeuGluTleAlaAaArgLeuGlnMet 224
Db 1812 AGCAAGTTCSSCAAGGCAAGCAACAACAACATGTTATCAACAAGTGCATCCAGCAACAAG 1871
Oy 225 TyrGluTleAaGlnHisProAlaValAaAaAaArgGluTyrHisTleAaHisValAla 244
Db 1872 CTCACAGCCAACTACATCCCCSSCAAAACAACAATGAAACCCAACTTTGGCACTC----- 1922
Oy 245 AlaAaen-----ThrGlyTleLeuValPhe-----GlnGlyPheThrTyrTleAaen--- 259
Db 1926 ---AACCATTTTTCAGAGCATGCTGTCAATATGAGACCAAGGCTTCTTGAGAAAGAACGA 1988
Oy 260 ---AlaPheAaenTTrpAlaValAlaValArgLeuSerPheTyrValGlyAaArgPheLeu--- 277
Db 1983 GACACCTTCATGGAGGAGACATTTATCCAGCTGTGTCACTCTCCAGGAAACAAGTTCTCAAG 2042
Oy 278 -----TleValLeuAaArgProAaAla 284
Db 2043 CAGATCTTCCAGGCGGAGATGTGCCATGGAGCGCCAGACCAAGAAACGCTGCCCACTT 2102
Oy 285 AaenSerAlaTyrGlnAaPThrTleGluTlePheLeuMetAlaSerArgAaPheCysAlaSer 304

```

Db	2103	AGCAGCCAGTTTCAACCGGCTACTGGAGCTGTGATATGGCCAGCGCTGGGTGCTGCCAGCC	2102
OY	305	PhePrryvi1eCyvAlGlnIeH1a1aPhePheAryLeuPheGluGluProLysePro	324
Db	2163	TTCTTTTGTCGATCATCAAGCCCAATGAGTTCT	2195
OY	325	LyseProLyseProValLeuPheSerArgLy-----SerSerPheHArgPheSerGly	341
Db	2196	---AAGAAGCCCACTGTCTTGCACCGGCACCTTGCGTGCAGCCAGCTGCCTGATCTCAGA	2252
OY	342	ArgHrghInLyse-----GlnValLeuAap	349
Db	2253	ATGATGGAGCCATCCGATCCGCCAGCTGGCTTACCCTACCTGATAGCTTGTAAG	2312
OY	350	TyrVal-----LysGluGly---354	
Db	2313	TTTGTGGAGCGGTACCGTGTCTGTGCAGGTGTGAAGCCCGGCTTCAACAGAGGGCAGC	2372
OY	354	-----354	
Db	2373	CTCCGCGGAACTTCCAGCGCATGCTGAGCTGTGCTGGCACCACGATAGCTGGCAG	2432
OY	355	---GlyHslyvi1eValGlnPheGluAryHsHsSerLy1eHsSerL1eArg-Se	373
Db	2433	ATGAGCAAAACCAAGATCTTTCTGAAAGACACACATGACATGCTGTGGAAAGTGAACCGG	2492
OY	373	rLeuAlaSerGlnProThrGluLeuAasSerGluValLeuGluInSerGlnInSerTh	393
Db	2493	GACAAAGCCATCACCGCACAGATCATCTCTTTCAGAAAGTCATCCGGGGATTCAAGAC	2552
OY	393	rSerLeuThrPhe-----GlyGluGlyAlaG1	402
Db	2553	AGGTCTAACTTTCTGAAGCTGAAGAGAGCGTGCACACTGATTCAGAGGCACTGGCGGGGT	2612
OY	402	uSerProGlyLy1eGlnSer-----CysArg-----410	
Db	2613	CACAACTGTAGAAAGACTACAGGCTGATGCGTGGGCTTCTCGCGGCTGCAGGCGCTG	2672
OY	411	---ArgGlyLyse-GluProLyseValSerAlaGlyGluProGlySerHs-----P	426
Db	2673	CACCGCTCCGGAAGCTGCACACAGTATCCGCTGGCCGCGCAGCGCATATCCAGTTC	2732
OY	426	rOserProAlaProAlaArgSer-----433	
Db	2733	CAGGCGCGTGCAGCGCTATCTGTGGCGAAGGCTTTCGCGCACGCGCTGTGGCTGTG	2792
OY	433	-----433	
Db	2793	CTCACCGTCAGGCTATGCCCCGGGCGATATGCGCCGCGAGGCTGCACACGCTCAG	2852
OY	434	-----ProAlaGlyAasLyvi1eGlnAlaPrrGlyAla-----Alas	445
Db	2853	GCTAGATATCTGTGGCGCTGTGAGGCTGTGAAGAAATGGCGCTGGCGGAGAAAGAACTT	2912
OY	445	erAlaProThrGluGluGluGluGluValValLyseAparGThrGlnInSerLyseProG	465
Db	2913	CGAAGAGAGATGAGCGCCAGAAAGAGCCAGAGAGGCGCGAGCGCATCAGAGAGCGC	2972
OY	465	LnProProGlnProSerThrGlySerLeuThrGlySerProHsLeuSerGluLeuSerV	485
Db	2973	CTGCGCCAGCTGGCTCGTAGAGACGTGAGCGGAGCT-----3010	
OY	485	alaAasSerGlnGlyV-----ValAlaProAlaAasValThreLeuSerP	500
Db	3011	--GAAGAGAAAGAGGCGCTCGGCGGAGAGAGAGGCTCTCGAGCGCATGAAAGGCGC	3068
OY	500	rOaenLeuSerProAapThrLyGlnAla--SerProLeuLeuSer-----514	
Db	3069	CG-----CATAGAGCTGTGATCATCACTCAACATGCTGCGACAAAGATTGTGGCTTC	3119
OY	515	-----ProLeuLeuAasPrrGlnAlaCyvProAryGThrAapPrr	527
Db	3120	CTGGGAGATTCAGGTGCTGCGCAGGCGCAGAGAGGCGCAGGCACTTATGTGCTTTGAGAGAC	3179

Qy	459	ThrgInGInserLyvProGInPrObroGInProSeThrGlySerLeuThGlySerPro	478
Db	3871	CTTCACCTGGAGGCTCTGCAGAGCAAGACGACGT-----TCTCCA	3912
Qy	479	-HiLeuSerGIuLeuSerValaAnSerGIuGLyValaLaProIaAnValaThLe	498
Db	3913	GGGCTCACTTAACAATAGACAGTACAGCCCTCTGGGACCCACTGGCGGGGACCGT	3972
Qy	498	uSerProAn-----LeuSerProAn-----Th	506
Db	3973	CTTCCCAAGAGGGGACTTTCACAGACTTCATCTGTCACAGCCCGACAGATGGAC	4032
Qy	506	rLyGInLaSerProLeuLLeSerProLeuLeuAnPrGIuLaCyvPro-----	523
Db	4033	CAAGCGATGAAGCCAGAGGAGCCGCTCTTGATTAATACAGACCCCTTCTGGAA	4092
Qy	524	-----AgThIraPArAGI	528
Db	4093	CTTTCTTCACCACTTAACCTGTGAATGAACACTTGCTTTTAAGACTAGACCTTT	4152
Qy	528	uAPrGIuGLyArGATG-----LyArGPrProThArLyArLyAlaLyrPhLeIaLy	546
Db	4153	GGACTCAAGAACAGAGAGCCCGCAAAAGATCCCAACACCA-----GTCTACG	4203
Qy	546	gGIuLaSerThrThrGIuArGThrThrLeuLyArLeuGIuValaIleThrSerTrpH	566
Db	4204	TGGCGCTTCAAGCTGGCAAGGACAGCTGCACACAGATCATGAACAACACTTAGTCA	4263
Qy	566	eGInSerThrValSerLyvGIuArLaImeProGIuLaLeuLyvSerLeuIlePheTr	586
Db	4264	CTGGGTGCACATCATGTGACCAAGAA-----AAGAGACAGACCTGCA	4305
Qy	586	oAnPrPheGIuLeuHILyArPheHILeThArPrPheLeuLyvGIuIleGIuInaLy	606
Db	4306	GGTGTTCACCACTGCGCAAA-----ACATGACT	4335
Qy	606	uAlaLeuTrpGIuGLyArGSetArvLaGLILeArGArTrpGIuInaLyArvA	626
Db	4336	GAGCAGCTTGAAGGCTGACGTCTATA-----GAATG	4368
Qy	626	ImeLeuLyArAnILeGIuGLyMeLyvHILeUaLaLanILeUaTrpLyvHILeSerGI	646
Db	4369	CATACCTTCCTCCCTGAGACATCCCTCACCAAGAGACATCTTAATTAAGGACACT	4428
Qy	646	uAlaLeuGLyLaLeuGLyLaAnGLyILeLyvSerSerArGTrpGIuAnPrPheCyAr	666
Db	4429	GGCC-----AAAGCTTGTCATCTGTGACTCTGTGT-----	4462
Qy	666	gAPrPheGIuLeuGLyValaCyvTrpLeuProLeuAnThrPheLeuLeuArGProLe	686
Db	4463	-----AGTTTACAAATAATGCTGT-----	4480
Qy	686	uHILeArGLyMeHILeTrpLyvGLInValaLeuGLyArGLyCyvLyvHILeProPhe	706
Db	4481	-----CTCTTAATCACTAAGATGTTACTACAG-----CACCAAA	4515
Qy	706	rHILeAlaArPheArGArCyv	713
Db	4516	ATATATTTCACTCAAGGCTGT	4537
RESULT 11			
US-10-559-415-189			
; Sequence 189, Application US/10559415			
; Publication No. US20060100132A1			
; GENERAL INFORMATION:			
; APPLICANT: AstraZeneca AB et al			
; TITLE OF INVENTION: Diagnostic Method			
; FILE REFERENCE: 101073-1P WO			
; CURRENT APPLICATION NUMBER: US/10/559,415			
; CURRENT FILING DATE: 2005-12-06			
; PRIOR APPLICATION NUMBER: 0313081.2			
; PRIOR FILING DATE: 2003-06-06			
; NUMBER OF SEQ ID NOS: 191			

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 189
; LENGTH: 7268
; TYPE: DNA
; ORGANISM: Homo sapiens - Base pairs nucleotide sequence
US-10-559-415-189

```

Alignment Scores:	
Pred. No.:	1.63
Score:	140.00
Percent Similarity:	35.1%
Best Local Similarity:	20.7%
Query Match:	2.6%
DB:	6
Gaps:	21
Length:	7258
Matches:	114
Conservative:	79
Mismatches:	202
Indels:	155
Gaps:	21

US-09-555-342B-2 (1-1045) x US-10-559-415-189 (1-7268)

405 glytyglinsertcyatgtagtglytylsgtupolysvalseralaglytuprogltyser 424
 3 gggccggagag-----cgccaaaggcgaagctgctctcaagctgcttggccaaaggagcg 56
 425 hisproserproalaProArgargserProalaglyAsnlyGlnalaaSpglYalaa 444
 57 ggaccacttcca-----gacctggggggcgccgtcgaaAaAaCGaAGCCTCA 104
 445 seralaprothngluIngluInglu-u-ValVallysaApargthngInglInserlyspr 464
 105 cctgctggccacttctactaactgaAGGCGTCTC----- 138
 464 oglnProProInProserthnglyserleuthrthgltyserProhlaeuserGluuse 484
 139 -ggcgccggcgagccggcggaaggcggggttctactaCaCGCTCTTCCaCCaGCC 197
 484 rValasnsrGlnnglyValalProalaAaValThrleuserProAsnleuserpr 504
 198 ctcaGactcAaAaAaCaGAGCTCTCTCaGAGTGTG----- 234
 504 aasprthlysglnalaserProleuileserProleuleuAaAspGlnalAcysProAr 524
 235 -ggcactcAGGAGGAGGCGCTGCCCaCaCCCTCTCTCaCTGACCaCaAGTaGTaATGa 293
 524 g-----ThraAspGlnAAspGlnlyAr 532
 294 GAAGGTGAGAACTCTCTCaTTCaGCTGGCTGaTgaTgaCCGGAGaAaAaCaGCTCCG 353
 532 gArglys--ArgPheProthAspLyshAlaTyRphelaaIalysgluValserThrh 551
 354 CAAGCGCTGGCGCTTGTGCTAGCaGCaGCGaGCGCTTTGaCaAAGAGCCCTTACaCaGCGCT 413
 551 rgluArghThyr-----LeuIyAspLeuGluVa 561
 414 GAATTCCTGACATGAGAGGCTGaAGATCCaGTGCGTGGAGCGaGCaATGTGGaACTTGGAGAG 473
 561 IilethrsertPheInserThr-----ValserlysgluAspAl 575
 474 cctcgAaAaCCaGAGCaCaCCaACaGCTTGAaAGGtGTGaAGAGGTGGCCaAGAGaCTGa 533
 575 aMetProGlnAlaLeuIyAserLeuIlePhe-----ProAsnPh 588
 534 CTTCACaCaCaCaCTCCaAGCGGCTCTGaGTaGCaCCaGACTCGaGCTGAAGATGaCGT 593
 588 egluProleuhtlysyRphelathAsnPhelueuIysguilegluIngluAla 608
 594 gGACaCTGCTGAGGCGGAGAAATGGGCaGCTGTGGCGAGCaAaCTCTCTCGaGCaGTC 653
 608 utrProgluIyArGserAsnAlaGlnIleArgAspTyrglnArgIleGlyAspValMetIe 628
 654 ATGGAGAGaCaTGAAGAGCGCTCCaGAGAGAGCaCaAAGAGaGATCGGaCTCGCTCGG 713
 628 ulysAsnIleGlnIyMetlyshleuAlaAlaniVleutRplyshleuserGluAla 648
 714 cCaGAGaGCaGAGGtGTGAAGCaCaAGCGGCTCA-----TCCGaAGTTCG 758

[illegible]

! ORGANISM: Homo sapiens	
US-10-505-928-450	
Alignment Scores:	
Pred. No.:	4.28
Score:	137.50
Percent Similarity:	33.1%
Best Local Similarity:	19.5%
Query Match:	2.5%
DB:	6
Gaps:	50
Length:	1195
Matches:	206
Conservative:	144
Mismatches:	378
Indels:	334

US-09-555-342B-2 (1-1045) X US-10-505-928-450 (1-11185)

[illegible]

```

Qy 293 GluPheUmeTalaSerArgAspPheCysUlySerPheTrpLysIleCysValGluHis 312
Db 7443 GAAGCAACACCTTACTGATGATTTCTGGCTAGAGCTTATGCTTT----- 7490
Qy 313 HisAlaPhePheArgLeuPheGluGluProLysProLysProValLeuPhe--- 331
Db 7491 ---GAATGGCCAAAGATTTGTTATCATCAGACCCAAACCATCTGACTGTATATGAA 7547
Qy 332 ---SerArgLysSer-----SerPheArgPheSerGlyArg 342
Db 7548 CTTTCTGGAGAAGATCTGGAGAACTGATATTGTTGATTCATTCATTCACATTCCTGCAACT 7607
Qy 343 Thr-----GlnLysGlnValLeuAspTyrValLysGluGlyHisValLys 358
Db 7608 ACTCAGGCAACACAGACAAAGAAAGACGACACCATTTGTTCTGATGGCTCCCTGAA--- 7664
Qy 359 ValGlnPheGluArgLysHisSerLysValLeuHisSerIleArgSerLeuAlaSerGln--- 377
Db 7665 ---AAACATCTGAGGTGCCAAGCGCTAAAGCTTAAAGCTTACTGCTGATGA 7709
Qy 378 ---ProThr-----GluLeuAsnSerGluValLeuGluGlnSerGln 391
Db 7710 TTCCCAACAGTTTCAAGTATGCTGCTGCTCTTCATTCAGACGAAACAAAGCTCCCTGAT 7769
Qy 392 SerThrSerLeuThrPheGlyGluGlyAla-----GluSerProGlyGlyGlnSerCys 409
Db 7770 CCACCTAGCAGACACTGCAAAATACAGTGTATATGAGAGCTCCACAGACGTAAGTTCCAA 7829
Qy 410 ArgArgLysGluProLysValSerAlaGlyGluProGlySerHisProSerProAla 429
Db 7830 GACGCTTACAGGAATTCAGAGATTCACCTTAAACCTAAC----- 7871
Qy 430 ProArgArgSerProAlaGlyAsnLysGlnAlaAspGlyAlaAlaSerAlaProThrGlu 449
Db 7872 ---AGAAAAAAACCCACTGAATAATATTATCATAGAC-----CTGGACACAA 7913
Qy 450 GluGluGluGluValValLysAspArgThrGlnGlnSerLysProGlnProProGlnPro 469
Db 7914 GAGACCAAGATTTAATATG----- 7934
Qy 470 SerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAsnSerGlnGly 489
Db 7935 ---ACAAATTACAGAGATGACATCTTGAATTTCTA----- 7967
Qy 490 GlyValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAspThrLysGlnAla 509
Db 7968 ---CCTGAGCTGACATCGGATAAATACTATCATCATGATATGATCATACT 8015
Qy 510 SerProLeuIleSerProLeuLeuAsnAspGlnAlaCysProAspThr----- 525
Db 8016 AATCCTGTGATGAAGACATCTTGGAAATGCAAACAGATATGATACAGAGTACCATCA 8075
Qy 526 ---AAspArgLysAspGlyGlyLysArgLysValArgPheProThrAspLysAlaTyr 542
Db 8076 GAACACATGACAGTAATGATGAAGTAATGATGACACACTCAAGTTCAGAGATGTCAT 8135
Qy 543 PheIleAlaLysGluValSerThrThrGluArgThrTyrLysLysValAspLeuGluValIle 562
Db 8136 GAGCAGCATGTCACCTTTCTTAACTGAGAAACATTTGAGGGCTCGCTGATGTTCTG 8195
Qy 563 ThrSerThrPheGlnSerThr-----ValSerLysGluAspAlaMetProLys 578
Db 8196 GCTAGCTACACTACGACAAACATGATATCATGACTTATGAAGATAG--AAGCCAACT 8254
Qy 579 AlaLeuLysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAspPhe 598
Db 8255 AGATCAACATGGGCTTCACTTCAACAATGGGATCCCTGC---TCTTACACAGAAACAGA 8311
Qy 599 LeuLysGluIleGluGlnArgLeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIle-Arg 618
Db 8312 ATTAGACCTTTT-----ACTTCCACAGGCAACATCCCT 8344
Qy 618 GAspTyrGlnArgLysGluAspValMetLeuLysAsnIleGlnGlyMetLysHisLeuAl 638

```

```

Db 8345 GCCAATTCCTCGTAAGTCTGCCACAGATATTCCAGAGATTGAAAGATTAATA----- 8396
Qy 638 AlaHisValLeuTrpHisSerGluAlaLeuGluAlaLeuGluAsnGlyLysSerSe 658
Db 8397 ---GCTGAAGCA---AAAGCCCTGATACATGTTTGAATCAG 8434
Qy 658 ArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnValCysTyrLeuProLe 678
Db 8435 CACTTGTCTGATGCT----- 8450
Qy 678 ValAsnThrPheLeuLeuArgProLeuHisValArgLeuMetHisTyrLysGlnValLeuGluArg 698
Db 8450 ----- 8450
Qy 698 GluCysLysHisValProProSerHisValaAspPheArgAspCysArgAlaAlaLeuAl 718
Db 8451 -----CAAGCTATTGC 8461
Qy 718 ArgLysIleThrGluMetValAlaGlnLeuHisGlyThrMetLysMetGluAsnPheG 738
Db 8462 AGACCAAAAGTAAATTAATACCA-----ACATGGGCCAATTGTAAGAGACTCA 8509
Qy 738 nLysLeuHisGlyLeuLysLysAspLeu-----IleG 749
Db 8510 GGAGAGATGATGAACAAABAAACATGCTGCTCTTCTTTTCAAGCAAGATTCCTTCAAG 8569
Qy 749 YLAspAsnLeuValValProGlyLysArgLysPheIleArgLeuGlySerLysLeuLysLe 769
Db 8570 AGCTAGAGAGGACTAGTACACATACCTCCCTATCTAAGTATTCTCTACCCAC----- 8624
Qy 769 uSerGlyLysGlyLeuGlnArgMetPhePheLeuPheAsnAspValLeu----- 786
Db 8625 ---CTTATGATCAGAGTGAACAGAGCTGCCTATGATGATGAAGATC 8671
Qy 787 ---LeuTyrThrSerArgGlyLeuThrAlaSerAsnGlnPheLysValHisG 803
Db 8672 CAATCCCCCATATTACACTGATACACATTGACAGTTTCAACATTTGGAAAGTGTCTTC 8721
Qy 803 YGlnLeuProLeuTyrGlyMetThrIleGluGlnSerGluAspGluTrpGly----- 820
Db 8732 TCAGACACATCATCTCCCTCACTATCTACTCAGGACGTGAACCTCTGGACACACAGA 8791
Qy 821 -ValProHisCysValLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAlaSerSe 840
Db 8792 GATCCCCCAGCCAGCTGCTGCCAGA-----ATAACGTCGGCTCATCTGT 8839
Qy 840 rArgSerGluMetGluLysTrpValGluAspIleGlnMetAlaIleAspLeuAlaGlyLys 860
Db 8840 AATGTCACACAGATTTCTTT---AAGAAATTCAATGAATATGTAAGCAACTTCA 8896
Qy 860 sSerSerSerProAlaProGluPheLeuAlaSerSerProAspAsnLysSerProAs 880
Db 8897 ACCATCAAGT-----GAGGAATACCTTCACATGAATGAGGCTCCCTTATATCTCCGA 8950
Qy 880 pGluAlaThrAlaAlaAspGlnGlnLysSerGluAspAsp----- 892
Db 8951 CACAAATTA-----GAACCTTGAGAAAGATGAAGTAAACCTGAGTTATTAAGA 9001
Qy 893 -LeuSerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgGlyAsnThrMetVal 912
Db 9002 AATGGAAGCTTCTCCACAGAACTT----- 9026
Qy 912 HisValCysTrpHisValArgAsnThrSerValSerMetValAspPheSerIleAlaValG 932
Db 9027 -----ATTGCTGTGGA 9037
Qy 932 ValAsnGlnLeuSerGlyAsnLeuLeuArgLysPheLysAsnSerAsnGlyTrpGlnLysLe 952
Db 9038 AGGAACCT-----GAGATTTCTCAAGATTTCCAAACAA--AACCAGATGCTCAAGTTT 9087
Qy 952 uTrpValValPheThrAsnPheCysValAspPheTyrLysSerHisGlnAspAsnHisAsp 972

```


Db 9088 CTGG-----AGAACATCAAGATG--TTTCC 9112
 Qy 972 oleuAlaSerIeuProIeuLeuGlyTyrSerLeuThrIleProSerGIuSerGIu 990
 Db 9113 CACCATTAACACCTGAGCTGGAACCTGTTATTACACTCCGATGAATTGAA 9167
 RESULT 13
 US-11-118-524-1
 / Sequence 1, Application US/11118524
 / Publication No. US20060088647A1
 / GENERAL INFORMATION:
 / APPLICANT: GU, WEI
 / TITLE OF INVENTION: ARF-BP1 AS MEDIATOR OF P53-DEPENDENT AND INDEPENDENT TUMOR
 / FILE REFERENCE: 19240-497US2
 / CURRENT APPLICATION NUMBER: US/11/118,524
 / CURRENT FILING DATE: 2005-04-29
 / PRIOR APPLICATION NUMBER: 60/610,506
 / NUMBER OF SEQ ID NOS: 22
 / SOFTWARE: PatentIn version 3.3
 / SEQ ID NO 1
 / LENGTH: 13128
 / TYPE: DNA
 / ORGANISM: homo sapiens
 / FEATURE:
 / NAME/KEY: misc feature
 / LOCATION: (3370)..(3372)
 / OTHER INFORMATION: CODES FOR A PROLINE AS COMPARED TO A LEUCINE AT AMINO ACID
 / FEATURE:
 / NAME/KEY: misc feature
 / LOCATION: (7918)..(7920)
 / OTHER INFORMATION: CODES FOR A LYSINE AS COMPARED TO A THREONINE AT AMINO ACID
 / OTHER INFORMATION: RESIDUE 2640 OF SEQ ID NO:2
 / FEATURE:
 / NAME/KEY: misc feature
 / LOCATION: (7924)..(7926)
 / OTHER INFORMATION: CODES FOR A VALINE AS COMPARED TO A SERINE AT AMINO ACID RESIDUE
 / OTHER INFORMATION: 2642 OF SEQ ID NO:2
 / FEATURE:
 / NAME/KEY: misc feature
 / LOCATION: (7927)..(7929)
 / OTHER INFORMATION: CODES FOR A Cysteine AS COMPARED TO A Serine AT AMINO ACID
 / OTHER INFORMATION: RESIDUE 2643 OF SEQ ID NO:2
 / FEATURE:
 / NAME/KEY: misc feature
 / LOCATION: (9955)..(9957)
 / OTHER INFORMATION: CODES FOR AN ASPARTIC ACID AS COMPARED TO A GLYCINE AT AMINO ACID
 / OTHER INFORMATION: RESIDUE 3319 OF SEQ ID NO:2
 / FEATURE:
 / NAME/KEY: misc feature
 / LOCATION: (11293)..(11295)
 / OTHER INFORMATION: CODES FOR AN ALANINE AS COMPARED TO A VALINE AT AMINO ACID
 / OTHER INFORMATION: RESIDUE 3765 OF SEQ ID NO:2
 / US-11-118-524-1
 Alignment Scores:
 Pred. No.: 5.72 Length: 13128
 Score: 137.00 Matches: 199
 Percent Similarity: 31.9% Conservative: 110
 Best Local Similarity: 20.5% Mismatches: 326
 Query Match: 2.5% Indels: 336
 DB: 7 Gaps: 47
 US-09-555-342B-2 (1-1045) x US-11-118-524-1 (1-13128)
 Qy 121 ProProAspHisThrGlnLeuGlnGluGlnLeuThrArgTyrLeuPheAlaLeuGlnVal 140

Db 9076 CTTCCA-----GCCATTGAGAGAGAGTACTGGCAGACGAGAGAGTGGACGAC 9126
 Qy 141 LyeGlnAspLeuAlaGlnGlyArgLeuThrCysAspAspThrSer----- 155
 Db 9127 CGAGAGAACTTACAGAGAAAT-----GCCAGCTCAGACACCCCTTATGACCTTGAC 9180
 Qy 156 ---AlaAlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAspGlu--- 173
 Db 9181 TTTCATCCAGACTGCGCCCTCAGACCTGCGCCGAGTGTCTCTAGAGGATATGAGGACAGT 9240
 Qy 174 -----AlaLeuAspArgGluHisIleValAlaLeuAspGlu--- 184
 Db 9241 GTGTAGCTGTGATGACCACTGACATTCGACGAGGCTCAAGCCCTGACAGCAGACCA 9300
 Qy 185 TyrIleProGlnIleAspAlaLeuGluAspArgIleValGluPheHisIleAspHis--- 203
 Db 9301 GAAGCCCGGACGACGACCTCATGACATGACGCTG-----TTTGGGACAGATGACACC 9354
 Qy 204 -----IleGlyGlnThrProAlaGluSer----- 211
 Db 9355 TCCGACCTCTGCTATTTCTCCGAGCCCGCTTTACACCGCTTAAGTGCACACCT 9414
 Qy 212 -----AspPheGlnLeuLeuGluIleAlaArg-----ArgLeuGluMetTyrGlyIle 227
 Db 9415 GGGATCCAGATATCTGCGCTTGTGTCAGAGAGTGGACACTTCCAGATGGGGGTATGC 9474
 Qy 228 ArgLeuHis-----ProAlaValAspArgGluGlyThrTyrHisIleAspHis----- 242
 Db 9475 AGCAGCATTAACAGGCTTGTGACATGATGATATCTCTCCGCTCCAGACGCG 9534
 Qy 243 -----AlaValAlaAspThrGlyIleLeuValPheGlnGlyPheThr 256
 Db 9535 CTCCTTGTGACACGACGAGAGCCCTTCTGTCTGTGTGCTGCTACTTTTGTGATGAGCCA 9594
 Qy 257 LysIleAspAlaPheAspThrAlaValAlaVal-----ArgLysLeuSerPhe-----LysArg 273
 Db 9595 AACCTCAATCTAGCCGTCTACACCCAGTACTGAGAAATCTGCTACCTGACCCAGACC 9654
 Qy 274 LysArgPheLeuIleLysLysLeuArgProAspAlaAspSerAlaTyrGlnAspThrLeuGlu 293
 Db 9655 GGCACCTGGGTCTACCTCC-----AGTGCCTCTCC 9684
 Qy 294 PheLeuMetAlaSerArgAspPheCysLysSerPheTrpLysIleCysValGlnHis 313
 Db 9685 ATCTTGAGCGGACGACGTAGACT-----GAGCTATGCAATGAACAACCC 9729
 Qy 314 AlaPhePheArgLeuPheGluGluProLysArgProLysProLysProValLeuPheSerArg 333
 Db 9730 AAACCTCACTCAAGTGAAGGAAAGGCAAAAGTGCAGCAGG-----AGCTGT 9777
 Qy 334 GlySerSer-----PheArgPheSerGlyArgThr 343
 Db 9778 GGGTCAGAGTCCATGAGAAACGTCCTCCCTGGACCTGTACACAAGATGAGTCAAAAGAGC 9837
 Qy 344 GlnLysGln----- 346
 Db 9838 TCCACACGCTTCTCGGCTCTCAGTATCCATGATGACGACCTTAGGCTGACGAGACTAAT 9897
 Qy 347 ValLeuAspTyrValLysGluGlyGlnLysLysValGlnPheGluArgHis--- 365
 Db 9898 ATATTTCAGATCCAGGCTTCCAGGGGGCGTAAACATACCGAG-----AAGCATGCA 9948
 Qy 366 -----SerValIleHisSerIleArgSerLeuAlaSerGlnProThrGluLeuAsn 382
 Db 9949 AGCGGTGACTCCACGTCACATTCATCCCAAGCTGCACCTGTGTCTGACACACGCTT 10008
 Qy 383 SerGluValLeuGluGln-----SerGlnGlnSerThr 393
 Db 10009 TTGGATATCACTCATTTGGCCAGAGTATTTCCAGACCATTCACACGACGCGGAGCC 10068
 Qy 394 SerLeuThrPheGlyGluLysGluSerProGlyGlyGlnSerCys----- 409

```

Db 10069 AAGAAACAACGTGTGAGTGTATGCGGGAAGGCAATAAGCCTGTAGCCATGCTCC 10128
Qy 409 ----- 409
Db 10129 TCACAGTCTCCACAGTGGCATTTGCACAGACTTCTGGGACTTATGGTAAACTGGAC 10188
Qy 410 -----ArgArgGlyLeuPro-----LysValSerAlaGly 420
Db 10189 AACATGATGTCCACCGGAAAGCAAGAACTCCGTAAGTCAAGTCCAGTACGCGCTGGC 10248
Qy 421 GluProGly----- 423
Db 10249 GGTAGGGGGAACCTCTCATACAGCTCGAGCCTCTCCACTGGGGCAGCTATGAAC 10308
Qy 424 -----SerHisPro----- 426
Db 10309 ATGTTGTACACCCAGTCATCCGCGGAGCTCTCTTAACTGAGAACTCCAGACTC 10368
Qy 427 -----SerProAlaProArgSerProAlaGlyAsnLysGln 439
Db 10369 CTTTCTCATCTCAATGCTCTCCAGAAAAGGTGTCAAGACAGGCTTAATCT 10428
Qy 440 AlaAspGlyAlaAlaSerAlaProThrGluGluGluGluValValLysAspArgThr 459
Db 10429 GGCAGCGGTGCTCTCCACACCACTGCCACTCAACCATCTACCAACACCACT 10488
Qy 460 GlnGlnSer---LysProGlnProProGlnProSerThrGlySerLeuThrGlySerPro 478
Db 10489 GCGCGCTCCACACGCGCCACACCCCTAATGACCCACCC---CCTGTCACTTCGCTCA 10545
Qy 479 HisLeuSerGluLeuSerValAsnSerGlnGlyGlyValAla----- 492
Db 10546 GCGCTGTGCTGCGCAGCGCTATTTCACCATGTGCTGCTGCTGCGACACAGTGACT 10605
Qy 493 ---ProAlaAsnValThrLeuSerProAsnLeuSerProAspThrLysGlnAlaSerPro 511
Db 10606 ACCCCACGACTGCTACACTACTGTTCAATTTCTTCCACTACTAAAGGACAGCAA--- 10662
Qy 512 LeuLeuSerProLeu---LeuAsnAspGlnAlaCysProAspGlyThrAspAspGlu 530
Db 10663 ---TCTCCAGGAAGTAGTAGTAGGGGACAGCAGTACAGAC--- 10704
Qy 531 GlyArgArgLysArgPheProThrAspLysAlaLysPheLeuAlaLysGluValSerThr 550
Db 10705 -----TTTAAAGATGGTGTCTCT 10722
Qy 551 ---ThrGluArgThrTyLeuLysAspLeuGluValLysThrSerTrpPheGlnSer 568
Db 10723 GGCCTCACTGAAACAGACTACAGCTCTGTAGAGGTGTGACATCCACTCTGTGTTCT 10782
Qy 569 ThrValSerLysGluAspAlaMetProGluAlaLeuLysSerLeuLysPheProAsnPhe 588
Db 10783 GAGGAGGCTTAAAGAGATGCA----- 10803
Qy 589 GluProLeuHisLysPheHisThrAsnPheLeuLysGluLeuGlnArgLeuAlaLeu 608
Db 10804 -----GCCAAGCTACTACTGCAGCTC----- 10824
Qy 609 TrpGluGlyArgSerAsnAlaGlnLysArgAspTyrglnArgLysGluAspValMetLeu 628
Db 10825 ---TCCGGGGGAGCTCGGAGCCCGGAGC-----ACTGTTCTCAAG 10863
Qy 629 LysAsnLysLeuGlnGlyMetLysHisLeuAlaAlaHisValLeuTrpLysHisSerGluAlaLeu 648
Db 10864 CTGCTACTGATGAGCCCGCATCTGGTTATACCTTTGTAAACAAATAGTACCTTG 10923
Qy 649 ---GluAlaLeuGluAsnGlyLysLysSerSerArgArgLeuGluAsnPheCysArg 666
Db 10924 CTGGCCAGGCTGCGGAAATACAACTCGACAGACAGCGGAGCCCA----- 10971
Qy 667 AspPheGluLeuGlnLysValCysTyrgLeu-ProLeuAsnThrPheLeuLeuArgProLe 686
Db 10972 -----TGTGAACCTCTCTCTGTAGTGGCTGAGTGAAGACAG 11010

```

```

Qy 686 WHisArgLeuMetHisIleTyrgLysGlnValLeuGluArgLeuLysHisHisProProse 706
Db 11011 CCACAGA-----CCACAGA 11025
Qy 706 rHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGlnLysThrGluMetValAlaGly 726
Db 11026 CTGAAG---GCCAAATGTCAGACAGCTTTG----- 11053
Qy 726 nLeuHisGlyThrMetLysMetGluAsnPheGlnLysLeuHisGluLeuLysLysVal 746
Db 11054 ---ACATGGCTGAGATG---TGGTAATTTGTGCGCATCTCAAGACGGA 11094
Qy 746 nLeuLysGlyLysLeuAsnLeuValProGlyArgGluPheLeuArgLeuGlySerLe 766
Db 11095 CTTTG---GCT---GSCCGGAG---CTCAGCTGCTCTTAT 11129
Qy 766 UserLysLeuSerGlyLysGlyLeuGlnGlnArgMetPhePheLeuPheAsnAspValLe 786
Db 11130 GTCCATGTTGACATCCAAAGACATCTACCCAGAACTTCTTGAAGGTACTACAGTCAAT 11189
Qy 786 nLeuTyrg-----ThrSerArgGlyLeuThrAlaSerAsnGlnPheLysValHis 802
Db 11190 CATCCAGCTCCGGGACGACAGCGCGG-----GCTAAACAAGAACCCAGACAGAC 11240
Qy 802 nGlyGlnLeuProLeuTyrgLysMet-----ThrLeuGlu 814
Db 11241 AGCGAGCTAGTGTCTCCGCTTTAAGCTCAGCTAAGACATCCAGCAGCTGCTCGCA 11300
Qy 814 UserGluAspGluTrpLysValProHisCysLeuThrLeuArgGlyGlnArg----- 831
Db 11301 GCTGAGGCTGAGGCTGATGCTCATTAACAATGTACGTGAGGTCAAAGGCGCGGAG 11360
Qy 832 ---GlnSerLysLeuValAlaAlaSerSerArgSerGluMetGlySerTrpValGluAs 850
Db 11361 ACACCAACAAGACAGCAAGCTGCGAGTCTACGACGTCAAGGCGCTGCGAGGAGGA 11420
Qy 850 pIleGlnMetAlaLeuAspLeuAlaGluLysSerSerSerProAlaProGluPheLeuAl 870
Db 11421 ATCACCCATGATGTGACCAAGCCATCTCCAGTGTCAAGATCAATCAATTCCTTC 11480
Qy 870 aserSerProPro-----AspAsnLysSerProAspGluAlaThrAl 884
Db 11481 CGATGAAACCCACAGGCGGAGGAAGAAAGAAAGAACCACTAGTTACCCCTGCT 11540
Qy 884 alaAspGlnGluSerGluAspAspLeu 893
Db 11541 CAGCGACAGCTGAGTTTGACGAGCTG 11568

RESULT 14
US-11-217-529-77055
; Sequence 77055, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: KODAMA, NORIHISA
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217, 529
; PRIOR FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77055
; LENGTH: 5040
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-77055

```

Alignment Scores:

Precl. No.:	1.75	Length:	5040
Score:	136.50	Matches:	180
Percent Similarity:	35.1%	Conservative:	145
Best Local Similarity:	19.5%	Mismatches:	366
Query Match:	2.5%	Indels:	235
		Gaps:	41

US-09-555-342b-2 (1-1045) x US-11-217-529-77055 (1-5040)

```

Qy 37 G1YLYSEUVALSER-----I1EYSE1LEGINMETEUEAPRTHING1U 52
    |||||
Db 1081 GGTAACTATTCTGTATCAAGTGTCTGAAAAGACACTGATGAAGAAAGATCA 1140
    |||||

Qy 53 A1APHEGLUVALPROG1NARG1A1PROG1LYLYEVALLEU1EUPR1A1VALCYSE1NH1S 72
    |||||
Db 1141 AAATTTCAATTACGAATTCAAATG----- 1164

Qy 73 LEU1EN1EUVALG1UG1YAP1TYRPHEG1LYEUG1UPHEPRO1APR1SLYS-----LYS 90
    |||||
Db 1165 -----GAGGATTTTACTAGAAATTA--GAGCATAAACCCCTGAA 1203

Qy 91 I1E1THRVAL1TRPLEUAPR1EUL1EYPR1O1LEVALYSG1N1E1ARG1PRO1YSH1S 110
    |||||
Db 1204 CTGATCTCTTTTAAAGAAAGACAAATTCATTGGAACATGAATAAACGTTCCACCGAA 1263

Qy 111 VALVAL-----VALLYSPHEVALVALYEPHEPHERPRO1APR1STH1NG1N1EUG1N 128
    |||||
Db 1264 CTACTAGAAACTGTTTCTTTGACTTAAAGCTTAAAGCAAGAAAGAAATTAACCTTCACTAGA 1323

Qy 129 G1UG1UL1EUTHR-----ARGTYR1EUPR1E1A1EUG1N1VALYSG1N-----ASP 143
    |||||
Db 1324 CAGAGATTAATGCTGTGAAGCAAAATATTCATTCTGTG--GTGAACAAACGTTGGAT 1380

Qy 144 LEU1A1--G1NG1YARG1EUTHRCYSE1N1APR1THSER1A1A1ALEU--LEU1LESER 161
    |||||
Db 1381 TTGGCCGTGCAAGTAAACTACTCTTTAAATACATCTGGATACAGAAACGGCATCG 1440

Qy 162 H1E1LEVALG1N1SERG1UL1EGLY1APR1E1APR1E1U1ALEU1APR1G1U1H1E1U1A1 181
    |||||
Db 1441 CCGTTATCGCAGAGTATGATATCATTAAGAAATTACTGCGAGCAGCAAAATATGCTGA 1500

Qy 182 LYSE1N1YSTR1E1PROG1N1ASP1A1LEUG1U1APR1Y1E1VAL1G1UPHE1NH1S 201
    |||||
Db 1501 AATGAAAT-----GATTCAGACCATCATTCGAAAGACTGCTGTAATTCAGTAAAT 1554

Qy 202 ---ASNH1S1LEGLYGIN1THRPRO1AG1USER1APR1EUG1N1EUG1N----- 217
    |||||
Db 1555 GTGAAATGAGCTA-----CAAGAAAAAATGTGTAACCTTGAAATTCGATTCGG 1602

Qy 218 ---I1E1A1ARG1EUG1U1METYR1G1Y----- 226
    |||||
Db 1603 ATTTTGCTGATTAATTGAATAATTATGAAGAAACAGACAGACGCTTACAAAAGTGA 1662

Qy 227 -----I1EARG1EUNH1SPRO1AL1Y1ASP 234
    |||||
Db 1663 GAAACACAGACCTTAAAGAAAGCAATTAATTCAGTCGAGAAACATCAATCGC 1722

Qy 235 ARG1G1UG1YTHRY1S1LE1N1E1U1A1VAL1A1ASN1H1R1G1Y1LE1U1VAL1PHE1NG1Y 254
    |||||
Db 1723 AAGATGAAACCCGATTAATATATCCT-ATTAAAGAGCGGATTCCTTACAAATTCGTGGC 1781

Qy 255 PHE-TH1LYS1LE1SN1A1PHE1ASNT1P1A1LY1VAL1ARG1Y1SE1USER1PHE1Y1ARG1Y 274
    |||||
Db 1782 ATCAACTGAAAGAAACAGGCCAACAAATTCGTCACTTCATGAGACGCCCTGGA 1841

Qy 274 SARPH1EUL1E1LYS1E1U1ARG1PRO1APR-----A1A1EN1SER1ATY 288
    |||||
Db 1842 AAAAGAAATTAAGAACTTGAGCGCAAACTTTCATCTTACAAATTCGAAATTCGTGCTAT 1901

Qy 288 RGIN1APR1THRYE-----G1UPHE1UE1CAL1SER1ARG1APR1PHECY1SY1SER1PHE1TR 306
    |||||

```

```

Db 1902 TATTCAGAAATTTAAGAAAGAGCTTGAATTTATTAAGAAATCAGAAATGCGATAA----- 1956

Qy 306 RLYSE1LECYSE1VALG1U1NH1E1N1E1A1PHE1PHE1ARG1EUPHE1G1U1PRO1YSPR 326
    |||||
Db 1957 -AAAACACTTTAAGAGATTTTGAATAATTTCAAGGCTTTGGCAAGAAAGAAAGAA----- 2010

Qy 326 OLYSPROVAL1EUPHESER1ARG1Y1SER1SER1PHE1ARG1PHE1SERG1Y1ARG1THRY1N1YSG1 346
    |||||
Db 2011 -----CGTATGTTAGGGA 2024

Qy 346 NVAL1EUPR1TYRVALY1G1UG1LYG1YH1SLY1LYS----- 358
    |||||
Db 2025 AGCAATAGACCCCTTAAAGCTGTAACCTGACAAACCAAAATCATGGCTTCATCATACAT 2084

Qy 359 ---VALG1N1PHE1G1U1ARG1Y1SE1SER1Y1S1E1S-----SER1LE1ARG1SE 373
    |||||
Db 2085 ACATGTTGAAAGAAAGAAAGAGCTCAACTGATACAGTCTAGAGATCAAAATCAAAATC 2144

Qy 373 R1EUA1ASERG1N1PRO1THRG1U1EUA1N1SERG1U1ALEUG1UG1N1SERG1N1SER1TH 393
    |||||
Db 2145 TTTGGAGTACAGAGATTTCCAAACTGAAAGAAAGAAAGAGCTTATATCCACGAAAGA 2204

Qy 393 R1SER1EUTHR1PHE1G1Y1G1UG1Y1A1AG1USER1PROG1Y1G1Y1N1SER1CYSE1ARG1ARG1Y 413
    |||||
Db 2205 ATCTTTTAACT-----AGAGATTTGCAACAAATGCTGCAGAAAGAAAAA 2246

Qy 413 SGI1UPR1OLY1VAL1SER1A1GLY1Q1UPR1OGLY1SER1H1SPRO1SER1PRO1A1PRO1ARG1SE 433
    |||||
Db 2247 AGAAATTCAGATAGCTTAAGAGAG-----TTGAAATTCAGTATATGAAATTAAT 2300

Qy 433 RPRO1AG1Y1ASN1Y1SG1N1A1APR1Y1A1A1ASER1A1PRO1THRG1UG1UG1UG1 453
    |||||
Db 2301 GGAATTTAGTTCAGAAAGAGGGCCAAATACAAAGCCAAATC-----AAAGAAATTCGA 2351

Qy 453 UVAL1VALY1SAPR1YTHRG1N1-----G1N1SER1Y1PROG1N1PRO1PROG1N1PROSE 470
    |||||
Db 2352 AATCAGCTTTGAAAGATTAAGTTCGTTCCGATTCATCAATCAAAAGATCAAA----- 2397

Qy 470 RTHRG1Y1SER1EUTHRG1Y1SER1PROH1S1EUSER1G1U1EUSER1A1N1SERG1N1G1Y 490
    |||||
Db 2398 -----GAAATGAAATCGATACGCTC 2417

Qy 490 YVAL1A1PRO1A1ASN1VAL1THRY1EUSER1PRO1A1N1EUSER1PRO1APR1THRY1S--G1N1 509
    |||||
Db 2418 ATGCAAGATTCGCAACTCAATGAGCTCAAAATACCATGATGATACAGAAATGAATAAT 2477

Qy 509 ASER1PRO1EUL1LESER1PRO1EUL1EUN1--ASRG1N1A1CYSPRO1ARG1THRY1SER1APR1 528
    |||||
Db 2478 GAACTCCTTAACCTGAATTAATCAATATAGAGAAACAAACAAATTAAGAAATTCATCAGA 2537

Qy 528 UAERG1UG1Y1ARG1Y1ARG1Y1APR1PHE1PRO1THRY1APR1Y1A1TYR1PHE-----I1 544
    |||||
Db 2538 AATAGAGAGCTTGATTAAGAACTAAGAAACCAAAATTCATATTAATTCCTTGATCA 2597

Qy 544 EAL1LYSG1U1VAL1SER1THRG1U1ARG1THRY1Y1EULY1APR1EUG1U1VAL1LE1THR1SE 564
    |||||
Db 2598 AATATGATGATGCAAGACCTGTAAGACCGACTTGAGAAAGAAATGAGCAAAATCCAAAGT 2657

Qy 564 R1TPR1PHE1GIN1SER1THVAL1SER1Y1SG1U1APR1A1ME1C1PROG1U1A1LEU1Y1SER1EUL1 584
    |||||
Db 2658 TCAACTGAAAGATGCAAACTCAGATACAGGCTTATGAGAGATTAATTAATTCCTCG----- 2712

Qy 584 EPHE1PRO1ASN1PHE1G1UPR1O1EUNH1S1YSPR1E1N1STH1ASN1PHE1EULY1EUG1UG1 604
    |||||
Db 2713 -----AATGAAATGCGCTGATGAATTCGAAATTA 2741

Qy 604 NARG1EUA1ALEUTHRG1UG1Y1ARG1SER1ASN1AG1N1LE1ARG1APR1Y1RGIN1ARG1LE1 624
    |||||
Db 2742 TGAACATAGCAAAACCAAGAAATTAATGATGCTTAATTAATGAACATGAGAAAGAG----- 2796

Qy 624 YAPR1VAL1MET1EULY1AS1N1LE1GIN1Y1MET1Y1S1E1U1A1A1A1H1S1EUL1TR1Y1SH1 644
    |||||
Db 2797 -----GAAAAATGGCGCAG 2810

```

[illegible][illegible]

```
Db      814 ACTGACCCCGAAGACAAAGTGAATTATGATTTAAGCAAGCCATGATGCATTA 873
Qy      718 AlagluIlethrGluMetValAlaGluLeuHisGlyThrMetIleLysMetGluAsnPhe 737
      874 AGGATTTTATGAAAAGATGTGATCAAGCAAGTGTCTGCA----- 915
      738 GlnLysLeuHisGluLeuLys-----LysAspLeuIleGlyIleAspAsnLeuValVal 755
      916 CAAAGATAGACATGATGTAAATTTGTGAACAAAGATTCCTTCAAAAAC----- 966
Qy      756 ProGlyArgGluPheIleArgLeuGly----- 764
      967 -----GAAATACGTAAATTTAGATTAATGACGAAAGAGAAAAATTAAACATGAA 1017
Qy      765 -----SerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArgMet 778
      1018 GGTATTAATCAACGAAAGAGTGTCTCAAAATCGAGCGAAACAGTTGTGTGACATACAA 1077
Qy      779 PhePheLeuPheAsnAspValLeuLeuTyr---ThrSerArgGlyLeuThrAlaSerAsn 797
      1078 TTTTATCTTCTAGATATATATGCTCTTATTTTAAAGGCCAAAGCAGTTAATAAATGCGAC 1137
Db      798 GlnPheLysValHisGlyGln-----LeuProLeu 807
      1138 CAGCATAGGATTCCTCAAGGCCCATTCATTCGCATTG 1176
```

Search completed: May 20, 2006, 04:15:02
Job time : 156 secs

QY 20 SerGlyIleSerThrLeuGluArgGlyGlnIys----- 30
DB 214 GCGCTCCGCCACCTGTGCGCCAGAGGACAGAGACATGAAATGATCCATTTGCGAGCT 273
QY 31 ProProThrProSerGlyIysLeuValSerIleValIleGlnMetLeuAspThr 50
DB 274 CCAGCTCAATCCCTGACGCGAGAG-ATGAGATGCACATCCGCGCTGCGAGACTCG 332
QY 51 GlnGluIlePheGluValProGlnArgAlaProGlyIysValLeuLeuAspAlaValCys 70
DB 333 GAGATCTCCGCCACATCCAGAGGAAACCAAGGCGATTCTTCATTTGACACATCTGCG 392
QY 71 AsnIleLeuAsnLeuValGlnGlyAspTyrPheGlyLeuGlnPheProAsnIleValCys 90
DB 393 AACCTACACAGCTGCGAGAGACTATTGGCATTCGCTGCGAGCCGAGAG 452
QY 91 IleThrValIlePheAspLeuLeuIysProIleValIysGlnIleArgAspProIleVal 110
DB 453 CAAGGCACTGCGCTTGAACCTAACAGTCCATCTTCAAGCAATGAAACTCATCCASCA 512
QY 111 ValValIalysPheValIalysPhePheProAsnIleSerIleGlnGlu 130
DB 513 TACACCATGCTGCTTAAAGTAAATCTTACCCACATGAACCTTGAAGATTAAAGAGAG 572
QY 131 LeuThrArgTyrLeuPheAlaLeuGlnValIysGlnAspLeuAlaGlnGlyArgLeuThr 150
DB 573 CTCACAGATACCTTTATACCTTCAGATTAAAGGACATTTTCATGCGCACTGCTG 632
QY 151 CysAsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGlnSerGlnIleGlyAsp 170
DB 633 TGCTCCCTTTCTGATGCTGCTTACCTGAGTCCGTGATTGTTCAAGCTGAGCTTGAT 692
QY 171 PheAspGluAlaLeuAspArgGlnHis-----LeuAlaIysAsnIysTyrIle 186
DB 693 TACGAT-----CCTAATGACATCTTGAATTAACATCACTGATGTTGAACTTTTC 743
QY 187 ProGlnGln--AspAlaLeuGlnAspIysIleValGlnPheHisAsnHisIle--- 204
DB 744 CCAGCAGCTGACAGAGCTGAGAAAGAAATATGGAATCTAAATAATGAACTCAGG 803
QY 205 GlyGlnThrProAlaGlnSerAspPheGlnLeuGlnIleAlaArgArgLeuGlnMet 224
DB 804 GGGGAGAGCCACAGCTGATTTAACTTGTCTCTGAAAGCTCACACTTGGAAACC 863
QY 225 TyrGlyIleArgLeuHisProAlaIysAspArgGlnGlyThrIysIleAsnLeuAlaVal 244
DB 864 TACGGGGGATCTCTCACCCCTGCAAGGATTCACAGGACACACATTTTATGAGTTC 923
QY 245 AlaAsnThrGlyIleLeuValPheGlnGlyPheThrIysIleAsnAlaPheAsnTyrAla 264
DB 924 ACAGCTCAGGCTTGTGTCTTTCAGGAAATAGAGAAATCCATTTGATTAATAATGCGCA 983
QY 265 LysValIArgIysLeuSerPheLeuValArgPheLeuIleIysIleValArgProAspAla 284
DB 984 GATGTCTGCAAAATTGAAGTTGAAGGAGACATTTATGATGTGACCCACAGAGAG 1043
QY 285 AsnSerAlaTyrGlnAspThrLeuGlnPheLeuMetIleAspArgAspPheCysIysSer 304
DB 1044 AAAAAAGCC-----ATGTTGGCATTTCAATCTTCAACACAGCTGCTGAAAAAT 1094
QY 305 PheTyrIysIleCysValGlnHisAlaPhePheArgLeuPheGlnIysProIysPro 324
DB 1095 CTTTGAAGGTGAGTGAAGAAACAGAGCTTTTATAGTATGCAAAATCCAGTCAATC 1154
QY 325 Lys-----ProIysProValLeuPheSerArgIysSerPheArgPheSerGlyArg 342
DB 1155 AAGCTGATATCAACACAGACATATTTTAAAGAGATGATTTTCATGATAGTGGGAAA 1214
QY 343 ThrGlnIysGlnValLeuAspTyrValIysGlnGlyGlnHisIysValIysGlnPheGln 362
DB 1215 GTTGCCAAAGAGGTGTGAG-----GCCAGTTCCAAAGATCCAG----- 1253
QY 363 ArgIysHisSerIysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu----- 380

DB 1254 AGGAGGCTCTGAGGTGACAGAGCAACATTTACTAGAGCGCGAGTTCCACTCTTG 1313
QY 381 -----LeuAsnSerGluValLeuGln-----GlnSerGln 390
DB 1314 AACAAACAGCTATCTTATACATGAAACCCCTGACGCCCTGCTTCTTCCCGAGGAG 1373
QY 391 GlnSerThrSerLeuThrPheGlyGlnGlyAlaGlnIysProGlyGlnSerCysArg 410
DB 1374 CAGAGAGAAACTTCTCTGAGTGAAGGTTCCT----- 1415
QY 411 ArgGlyIysGlnProIysValSerAlaGlnIysProGlySerHisProSerProAlaPro 430
DB 1416 -----AAAGAGAGAAACATTTCTGCTCCCTGATGCTCCAGCTCCCAAGAGGAGGCC 1469
QY 431 ArgArgSerProAlaGlnIysGlnAlaAspGlyAlaAlaSerAlaProThrGlnIys 450
DB 1470 CCGGAG-----TATGAGATCCCTGATGAA 1496
QY 451 GlnGlnIysValIalysAspArgThrGlnGlnSerIysProGlnProProGlnProSer 470
DB 1497 GAGAGATTAATAATAAGAA-----GAACCTTTA 1526
QY 471 ThrGlySerLeuThrGlySerProHisLeuSerGlnLeuSerValAsnSerGlnGly 490
DB 1527 ACC-----ATCTGAACTAGTATCAACCCAGAGTGCAGC 1562
QY 491 ValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAspThrIysGlnAlaSer 510
DB 1563 CTGCTCCCAACCCCTGT-----GATGACGATGAGATTGAC 1598
QY 511 ProLeuIleSerProLeuLeuAsnArgAlaCysPro-----ArgThrAspAsp 527
DB 1599 ATGCTCTTGAC-----TGCTCTTCTAGCGCTTGAGTGAAGA 1637
QY 528 GlnAspGlnIysArgArgIysArgPheProThrAspIys--AlaTyrPheIleAlaIys 546
DB 1638 GAAGACACAGATTCATTTGAGATCTGAAGACAGATGAAACCCCTTTGATGATGCGAA 1697
QY 547 GlnValSerThrThrGlnArgThrTyrLeuIysAspLeuGlnValIleThrSerThrPhe 566
DB 1698 GAAGAGAG-----CTGAAGAGGCTGCGCTGCTTGTGTGAGC 1739
QY 567 GlnSerThrValSerIysGlnAspAlaMetProGlnAlaLeuIysSer----- 582
DB 1740 TAGACATTTCTGACTGCGCATTTCCGGTGAACCACTGCTCAAGATTTTCCAGGCTC 1799
QY 583 -----LeuIlePheProAsnPheGlnProLeuHisIys 593
DB 1800 CTTGTGGTGGGCTGGGACTGCTGCTTTGTGATTTCCCTGCTCTCTCTCTTTTGAG 1859
QY 594 -----PheHisThrAsnPheLeuIysGlnIleGln 604
DB 1860 TCAGTATTGATCTCTCTTTATGCGAAATCGCGAG 1898

RESULT 2
US-11-412-475-7
; Sequence 7, Application US/11412475
; GENERAL INFORMATION:
; APPLICANT: The Salk Institute for Biological Studies
; APPLICANT: Novacta, Ifrls
; APPLICANT: Verma, Inder M.
; APPLICANT: Lockhart, David J.
; APPLICANT: Barlow, Carolee
; TITLE OF INVENTION: ANXIETY RELATED FACTORS AND METHODS OF USE
; FILE REFERENCE: 7518-70020-02
; CURRENT APPLICATION NUMBER: US/11/412,475
; PRIOR APPLICATION NUMBER: US 60/675,604
; PRIOR FILING DATE: 2005-04-27
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7

LENGTH: 2332
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (275) .. (2026)
US-11-412-475-7

Alignment Scores:
Pred. No.: 1.4e-24 Length: 2332
Score: 535.00 Matches: 197
Percent Similarity: 40.7% Conservative: 133
Best Local Similarity: 24.3% Mismatches: 288
Query Match: 9.8% Indels: 193
DB: Gaps: 30

US-09-555-342b-2 (1-1045) x US-11-412-475-7 (1-2332)

Qy 7 ArgProThrProGlySer-----ArgLeuGluAlaProGluAAsnSerGlyIleSer 23
Db 32 CGCCCAAGTTCGCGCGCGCTTCCCGCGCGCTCCCGCGCGCTGACTTCCCGCGCTTTC 91
Qy 24 ThrLeuGluArg-----GlyGlnIleProProProPro 35
Db 92 CGCCTTCTGCTCCCGCGCGCGCGGCTTCCGCGCGCGCTCCCGCGCGCGCGCG 151
Qy 36 Ser----- 36
Db 152 AGCTGCTGCTCACTCCGCCAGCGCGCGGCAATTAATATGATCGGCTTGAGTG 211
Qy 37 -----GlyLeuValSer----- 41
Db 212 GAGGGGCGCGCTGCGCGCTGCTGCTTCTCTCTGCTGCGAGACACACACACC 271
Qy 42 -----IleValIleGlnMetLeuAsp 48
Db 272 GACATGGGCTGTTTCTGCGCTGTTCCGGAGAATTTTACTCGAAGTTTCTCTCGGAT 331
Qy 49 AsnThrGlnGluAlaPheGluValProGlnArg-----AlaProGlyLeu 63
Db 332 GAATCCCAAGTTAACCTTTACACCCAGCAGCGGCGCATCAAGAACTCAACGAAAGTTCC 391
Qy 64 ValLeuLeuAspAlaValCysAsnIleLeuAsnIleGluIleAspTyrPheGlyLeu 83
Db 392 GTTCTCTTGACAGATTCCTACACGTAACCTTTGGAGATATATTTGGGCTA 451
Qy 84 GluPheProAsnIleValIleThrValTyrLeuAspLeuLeuValProIleValIle 103
Db 452 CGTTACTGTACAGAAAGCATCAGATGATGGCTGATCTGGCAAAACCTTGCTGAA 511
Qy 104 Gln-----IleArgArgProIleValValIleValIleValIlePhePhe 120
Db 512 CACAAGAACTGATCAACATGACCTTCATATCTTGTATTTGGTATTAATTTCTAT 571
Qy 121 ProProAsnIleThrGlnLeuGlnGluIleuLeuThrArgTyrLeuPheAlaLeuGlnVal 140
Db 572 GCTGAAGTCCATGATTAAGAAAGAAATTAACACATATCACTTTTCTTGAGGTG 631
Qy 141 LysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAlaIleLeuIle 160
Db 632 AAGCAAGATGCTCTTACGGCGCGCTGCTCCCGTCAACACGCTGCTCAGTGGGA 691
Qy 161 SerIleValIleGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgGlu---His 179
Db 692 GCGATGCGCATCCAGTCCGAGCTGGAGATTAAGCCATATAACATATGCAAGAT 751
Qy 180 LeuAlaValAsnIleValPro---GlnIleAspAlaLeuGluAspIleValGln 198
Db 752 GTATCTGAGTACCGGCTTGTCTGATCAAGAAAGAACTTGAAGAAAGCCATTAAGG 811
Qy 199 PheIleValAsnIleValIleGlnThrProAlaGlnSerAspPheGlnLeuGlnIle 218
Db 812 ATTCAATAAAGCTTAATGAGTCAATCTTCTGAGCTGAGCTGAATTACTTACGAGCT 871

Qy 219 AlaArgArgLeuGluMetTyrGlyIleArgLeuHisProAlaValAspArgGluGlyThr 238
Db 872 GCCAAATCCCTGGAGATGATGCGTGCATCTCACTCCGCTATGAGAAACAACTCT 931
Qy 239 LysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnIlePheThrIle 258
Db 932 GAGTATTTCTTAGATTAACCTCCGTTGCTGTTGTGTCAGAAATTAAGCAAGTG 991
Qy 259 AsnAlaPheAsnThrAlaValArgIleLeuSerPheIleValGlyValArgPheLeuIle 278
Db 992 GGGAAAGTATTTCCGCTCCGATTAACAAGGTTCACTTCAAGGAGCTCAATTTGAATC 1051
Qy 279 Lys---LeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAla 297
Db 1052 AGAGTACTGGGAAAGATTTGATAC-----GAACTCATTTCTTTTGAAGCTCGG 1102
Qy 298 SerArgAspPheCysIleSerPheTyrIleCysValGluHisAlaPhePheArg 317
Db 1103 AGTAAACTGCTTCAAGACCTCTGGAAGTGCAGTGGAAATCATATCACTTTTATGA 1162
Qy 318 LeuPheGluGluProIleProIleProIleProValLeuPheSerArgGlySer----- 335
Db 1163 ATGCCAGAAATGAATCCATTCATGTCAGAAAGAACTCAGCAAGTTTGATCCATAGT 1222
Qy 336 ---SerPheArgPheSerGlyArgThrGlnIleGlnVal---LeuAspTyrValIleGln 353
Db 1223 TATTAAGCACCGCTACAGTGGAGGACAGCTTTGCCAAATGACCGAGATCTTTCTATTCAG 1282
Qy 354 GlyGlnIleIleValGlnPheGluArgIleHisIleSerIleHisIleSerIleArgSer 373
Db 1283 CTTCGCCGCGCTGATCAAGATGTGACAAAGTGAAAGCAAGATTCACCTTAAGCGATA 1342
Qy 374 LeuAlaSerGlnProThrGluLeuAsnSerGluValLeuGlnIleGlnIleSerThr 393
Db 1343 GCACAAACACAGCGCTGATCAACACCC-----ATCAGTGAAGTAACTGCA 1390
Qy 394 SerLeuThrPheGlyGluGlyAlaGlnSerProGlyGlnIleSerCysArgArgGlyLeu 413
Db 1391 AACATG-----GAAATGCAAGAAATGAAGAA----- 1417
Qy 414 GluProIleValSerAlaGlnIleProGlySerHisIleProSerProAlaProArgArgSer 433
Db 1418 ACATTAATAATTTATGCA-----CCTTCAACGAAATAAGCTTTAAG 1459
Qy 434 ProAlaGlnIleValGlnIleAspGlyAlaAlaSerAlaProThrGluGluGln 453
Db 1460 AAAGCAAGATGAATAATAGCCCTGAT----- 1486
Qy 454 ValValIleAspArgThrGlnIleSerIleProGlnProGlnProSerThrGlySer 473
Db 1487 -----ACCCAAAGAAAGAAATCTCATGACCGTGGGAAAGAAATGGCCCC 1531
Qy 474 LeuThrGlySerProHisLeuSerGluLeuSerValAsnSerGlnIleValAlaPro 493
Db 1532 CAGAGTGAAGCTTACAT-----TCTCCC 1555
Qy 494 AlaAsnValThrLeuSerProAsnLeuSerProAspThrIleGlnIleSerPro----- 511
Db 1556 AGTATGCACTAATGCGCAAAAGTTC---CCTTACACGCGCTCCGCAAAACCTCTCTGT 1612
Qy 512 -----LeuIleSerProLeuLeuAsnArgIleAlaCysProArgThrAsp 526
Db 1613 GGAAGTGAATGATTTCTGTACACCTGTGTGACGAGAGAAAGCCATTAACAGTGGTGA 1672
Qy 527 AspGluAspGluIleArgArgIleValArgPheProThrAspIleValTyrPheIleAlaIle 546
Db 1673 GATTCAGATCTTAAGAAAGAGAGGATCACTTACAGCTGTAAACACAGAGTGTAGT 1732
Qy 547 GluValSerThrThrGluValGlyThrIleLeuIleAspLeuGluValIleThrSerTyrPhe 566
Db 1733 GAATCAGAAATTTCTAATAGAAACACCGGAAAG----- 1768

QY	656	GINSERTVALISLSELYGLIAAPRLAMETPROGIIWALALEULYSETEULIEPHERO	586
Db	1769	AGAAACAGAAATACGGCAGAGAAATATATGTGTGTTACGCC-----CCT	1813
QY	587	ASPHEGULPROLEUHHLEYPHEHISTHRAAPHEULYGLULGLUGLARGLEU	606
Db	1814	CAGTGGAGAGGTGA-----TTAAGGACACAAAGAAAAAACC	1852
QY	607	ALALEUTRGLUGLYARGSERASNALGINLEARGAPRYRGLARGLLEGLYASPRAL	626
Db	1853	CAAGCCGACCCCAACAGCAGCGGATCCACACACAGATCTCGTGGAGAAAGCCCGATATTC	1912
QY	627	METLEULYSAENILLEGINGLYMETLYSHLEUALALAHISLEUTRPLYSHISERGLU	646
Db	1913	CAAGCAAAAGAA-----GAGTATYGACACACACATTCAA	1945
QY	647	ALALEUGLWALALEUGLWANGLYTLELYSSERSEARGARGLEUGLWAPHECYARG	666
Db	1946	AAAGAACTTGTGGATCCATCCGAGATGTCCGAGAA-----	1981
QY	667	ASPHEGULEUGLWALYVALCYETRYLEUPROLEUANTHRPHELEULEARGPROLEU	686
Db	1982	CAATTAAAGAGATTCATCATTAATAAATAGAGAAC-----ACAAGGTGACCC	2029
QY	687	HIS-ARGLEUMETHISTRYRGLYSGINALLEUGLWARGLEUCYULYSHISHSIPROFSE	706
Db	2030	AATCGCATCCAGGCAT-----TCTCATTCGCCACGGAAG	2062
QY	706	RHLEIALASPHEARG-----ASPCTSE-----ARGALAALEUALAGLU1	720
Db	2063	TTACCCGACAGATGCGCAGTCCCACTGTTCAGATGCGGAGGAGTACAGTCTCTCGGAAGT	2122
QY	720	e-----ThrglMetValAlaGlnLeu	727
Db	2123	GAATTCAAAAACAGATCTTTGTAACCAACCACTT	2153

```

RESULT 3
US-11-412-475-5
: Sequence 5, Application US/11412475
: GENERAL INFORMATION:
: APPLICANT: The Salk Institute for Biological Studies
: APPLICANT: Hovatta, Ilirix
: APPLICANT: Verma, Indar M.
: APPLICANT: Lockhart, David J.
: APPLICANT: Barlow, Carolee
: TITLE OR INVENTION: ANXIETY RELATED FACTORS AND METHODS OF USE
: FILE REFERENCE: 7518-70020-02/US/11/412,475
: CURRENT APPLICATION NUMBER: US/11/412,475
: PRIOR FILING DATE: 2006-04-26
: PRIOR APPLICATION NUMBER: US 60/675,604
: PRIOR FILING DATE: 2005-04-27
: NUMBER OF SEQ ID NOS: 115
: SOFTWARE: PatentIn version 3.3
: SEQ ID NO 5
: LENGTH: 2487
: TYPE: DNA
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (52)..(1716)
US-11-412-475-5

Alignment Scores:
Pred. No.:          5.64e-22          Length:          2487
Score:              492.50             Matches:          163
Percent Similarity: 45.5%             Conservative:     99
Best Local Similarity: 28.3%           Mismatches:      216
Query Match:        9.0%               Gaps:            21
DB:
US-09-555-342B-2 (1-1045) x US-11-412-475-5 (1-2487)

```

43 LysIleGlnMetLeuAspAspThrGlnGluAlaPheGluValProGlnArg----- 59

Db 91 GAAAGTTTGGCTCTTGAAGAGTCCCAAGTCACTCCCTCAACCCCAAGCAAGGAGCATCAAG 150
 Oy 60 -----AlaProGluYLyValLeuLeuAra1aValCysAaenhiVLeuAenLeuValGlu 77
 Db 151 AAGTCAACAAAGAGGCTGTGTGTCTTGTCTTGAACACGGTTTTCCGGTCAATAACTCCGTGGAG 210
 Oy 78 GlyAerPTrPheGluPLeuGluPheProAerhiVsuVsuV1LeThValTrPLeuAerLeu 97
 Db 211 ATAGATTATTTTGGGCTGGGCTAGCTGTGACGAAGCCATCAAGACATATATGGTGGATCCC 270
 Oy 98 LeuYsProIleValLyseGln-----LLeaGrAerProLyenhiValValValLy 114
 Db 271 GCAAAAGACCTTGGCAAGACAAAGAGGCTGATCAACAGTCCGATTAATCTTTGAT 330
 Oy 115 PheValValLyserPhePheProAerhiVThrGlnLeuGlnGluLeuThraGTr 134
 Db 331 TTGGTATTAATTTCTATGCTGAAGATCCATGTAACTCAAGAAATAACTCAAGATAT 390
 Oy 135 LeuPheAlaLeuGlnValLyseGlnAerLeuAlaGlnGlyAerLeuThCysAaAerThr 154
 Db 391 CAGTTTTTCTTACAGGTGAACAAAGATGCCCTTCAAGGGCCGGCTCCCTGGCTCAAT 450
 Oy 155 SerAlaAlaLeuLeuIleSerhiVLeValGlnserGluIleGlyAerPheAerPdu1a 174
 Db 451 ATGTGCTCAAGATGGGGGCAATGATGCCATCCAGGGCGAATGGGAGACACGACCCATAC 510
 Oy 175 LeuAerAerGlu-----hiLeuAlaLyAenLyTrIlePro-----GlnGlnAra1aLeu 192
 Db 511 AAGCAACATGGGGGGTATGTGTCCGAGTACCGGGTTGTCTTGATCAAGAGAGAGCTG 570
 Oy 193 GluAerLyV1leValGluPhehiVhiVhiVhiV1leGlyGlnThrProAlaGluSerAer 212
 Db 571 GAAGAACCAATAGAAAGGATTCATTAATCACTTAATGGGTCAAGGCTCTTCCGAAGCTAG 630
 Oy 213 PheGlnLeuLeuGluIleAlaGrAerGlnuLeuMetTrGly1leAerLeuhiVProAla 232
 Db 631 CTGAATTACTTGAAGAGTCCCAATGCCCTGGAGAGTATGGTGTGAGCTCCATCTGTTC 690
 Oy 233 LyuAerAerGluGluThLyV1AenLyLeuAlaAlaAenThGly1leLeuValPhe 252
 Db 691 TATGGAGAAATAGTCCGAGTACTTCTTGAAGCTTGAAGCTCCATCGGAGTGTCTGAT 750
 Oy 253 GlnGluPheThLyV1leAenAlaPheAenTrPAlaLyValAerGlyLeuSerPheLyV 272
 Db 751 AAGAATAAAAGCAAGTGGGAGAGATTTCTGGCTCGGATTAACAAAGTCACTTCAAG 810
 Oy 273 ArgLyAerPheLeuIleLyV-----LeuAerProAerAlaAenSerAlaTyGlnAerThr 291
 Db 811 GAAACCCAGTTTGAAGTCAAGACTGTGGGAAAGACTGTAT-----GAAACCTCA 861
 Oy 292 LeuGluPheLeuMetAlaSerAerAerPheCysLyVserPheTrPdu1leCysValGlu 311
 Db 862 TTTCTTTTGAAGCTCAAGCAAAAGCTCTTGAAGCACTCTGGAATGAGAGCTAAG 921
 Oy 312 hiVhiSAlaPheAerAerGlyLeuPheGluGluProLyVProLyVProLyVProValLeuPhe 331
 Db 922 CACCATAGCTTTTCAAGATCCCGGACAGATCAATCAATTCATTATCAAGAAACTCAGC 981
 Oy 332 SerAerGlySer---SerPheAerPheSerGlyAerThGlnLyV1leGlnVal---LeuAer 349
 Db 982 AAGTTTGGGCTCAATAGTTATTAACATCGGTACAGGACAGCTTTGCAATGAGCGCAAT 1041
 Oy 350 TyrAllyVsuGluGlyhiVhiVhiVhiV1leGlnPheGlyAerGlyVhiSerLyV1lehiV 369
 Db 1042 CTTTCTATCACTTCCCGGCGCAATCAAGACGCTGGAAGAGTCAAGCAAGCTTAC 1101
 Oy 370 Ser1leAerSer1leuAlaSerGlnProThGlyLeuAenSerGluValLeuGlnInser 389
 Db 1102 CCCAAGAGGATGACACAGACTCAAGCTCAAGCTCGATCAAAAC-----ATCAATCGATTA 1155
 Oy 390 GlnGlnserThrSerLeuThrPheGlyGlnGly-----AlaGluSerPro 404

```

Db      1156 ACTGCAACACGAAACCGGAGATAGAGAAACCAAGATCATTCGCTCCCA 1215
Qy      405 GlyGlnSerCyArgArgGlyValSerValSerIaGluProGlySer 424
Db      1216 ---GTGAAAGTTTAAAGAAAGCAAG-----AAT 1242
Qy      425 HisProSerProIaIaProArgArgSerProIaGlyAenlyGlnIaIaSerGlyValIaIa 444
Db      1243 GAAGAACGCTTGACCCCAAGAAC-----AAATCT 1275
Qy      445 SerIaIaProIaGluGluGluGluValIaIaSerArgThr-----459
Db      1276 CATGACCTTGGAAGAAACGAGCGAGTGACTATACACTCTCCAGTACCGC 1335
Qy      460 -----GlnIaIaSerIaProIaIaProIaIaProIaIaSerIaIaSerIaIa 476
Db      1336 ACTAGTACCAAGATTTCTCTGTCGCGAGCGCAACTCTCTGCGAGCGACAAAT 1395
Qy      476 YSerProIaIa-----LeuSerGluLeuSerValAs 486
Db      1396 GACTCTTCAACAATGATGAGCGAGAAAGCTCAACAAGTGGGAGAGACTCAGATCTC 1455
Qy      486 nSerGlnGlyValIaIaProIaIaSerValIaIaSerProIaIaSerProIaIaSer 506
Db      1456 AAGCAGAGAGAGAGTCAAGCTCAGCTGTAACA-----1489
Qy      506 rLyGlnIaIaSerProIaIaSerProIaIaSerProIaIaSerProIaIaSer 526
Db      1490 ---CAAGCAGTGTAGTCAAGAAATCTTACA-----1522
Qy      526 PArgGluAerGluArgArgGlyValArgPheProIaIaSerValIaIaSerValIaIa 546
Db      1523 ---GAGAACACCGGAAAGAAAGAAACGAAACCGCCAGAGAAATGATGATGTTACCTG 1578
Qy      546 gLysIaIaSer-----ThrThrGluArgThrTyLe 556
Db      1579 GGGCTCAGTGGGAACCGGTTAAGAGAGAGAAAGAAAGAAAGCAAGCGGACCTTAC 1638
Qy      556 u-----LysAerPhe-----GluValIleThrSer 564
Db      1639 AACCGAGATCCAGACACAGATCTGCTCAAGAAATCTGACATCC 1684

RESULT 4
US-10-448-162A-1
; Sequence 1, Application US/10448162A
; GENERAL INFORMATION:
; APPLICANT: Hart, Matthew J.
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES RELATED TO A GUANINE
; FILE REFERENCE: ONYX 1023.DIV2
; CURRENT APPLICATION NUMBER: US/10/448.162A
; CURRENT FILING DATE: 2003-05-27
; PRIOR APPLICATION NUMBER: 09/865,960
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 08/943,768
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 60/029,979
; PRIOR FILING DATE: 1996-11-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 3150
; TYPE: DNA
; ORGANISM: Human p15 GEF-Rho
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (55) .. (2790)
; US-10-448-162A-1

Alignment Scores:
Pred. No.: 0.00555 Length: 3150
Score: 180.50 Matches: 187
Percent Similarity: 34.6% Conservative: 125

```

```

Best Local Similarity: 20.8% Mismatches: 351
Query Match: 3.3% Indels: 239
DB: 6 Gaps: 35

US-09-555-342B-2 (1-1045) x US-10-448-162A-1 (1-3150)
Qy      142 GlnAerPheAaIaGlnGlyArgLeuThrCyAenAerThrSerIaIaIaLeuIleSer 161
Db      58 GAAGCTTCCGCTCCAGAGGCGGCTCTCCAGAGCCCTCCGAGCTGCTGTTCCGCTC 117
Qy      162 HisIleValGlnSerGluIleGlyAerPheAerGluIaIaLeuAerArgGluHisIleVal 181
Db      118 AGCATTCATGGGCTGAGATGAGATGATTTTGAACAAGCTGAGACAAACTCAGAAAG 177
Qy      182 LysAenlyTyIleProGlnGlnAerPheAaIaLeuGluAerlyIleValGluPheHis 201
Db      178 CAAGAAC-----183
Qy      202 AsnHisIleGlyGlnThrProIaGluSerAerPheGlnLeuGluIleAaArg 221
Db      184 -----AGCCAGTTCCAGAGCTGAGACAGTGAAGCG 216
Qy      222 -----LeuGluMetTyGlyIleArgLeuHisAerProIaIaAerArgGluGlyThrLys 239
Db      217 CGCCAGCCCACTCATAGGCTCCCTGACGACAGTGGCC-----255
Qy      240 IleAenLeuAaIaIaAenThrGlyIleLeuValPheGlnGlyPheThrLysIleAen 259
Db      255 -----255
Qy      260 AlaPheAenThrAlaLysValArgLysLeuSerPheLysArgPheLeuIleLys 279
Db      256 -----CTGCAAGTTTGAAGCCAGAGACCCCTGCTTGCCTG 288
Qy      280 LeuArgProAerPheAaIaAenSerIaIaTyGlnAerThrLeuGluPheLeuMetAaIaSerArg 299
Db      289 CTGCATGCCAGCATGCTGGCTCACTG--GCCCCAAGAGGCGCAAGAGCCCTTCTCTG 345
Qy      300 AserPheCyLysSerPheThrPheIleCyValGluHisIaIaPhePheArgLeuPhe 319
Db      346 GACTTTCACCAACACTTC-----CTGAGAGAGACAGGCGTTCCTCCGGGTG-- 390
Qy      320 GluIaIaProLysProLysProLysProValLeuPhe-----SerArgGlySerSerPhe 337
Db      391 -----CCGGTCCCTCCCAAGCTGCTTGAATCTTGAACCGCATAGAGGCTGAC 438
Qy      338 ArgPheSerGlyArgThrGlnLysGlnValLeuAerTyValLysGluGlyHisLys 357
Db      439 CTGATCTCCAGAGATGTCAGCGCGGTGTCGAGAGGATGATGCAAGC-----CAG 492
Qy      358 LysValGlnPheGluArgLysHisSerLysIleHisSerIleAerSerLeuAaIaSerGln 377
Db      493 CAGGTACCGCTGGGCGGAGCTGAGAGACTTCGTTCCAAAGCGGCTCATAGGCGATGACG 552
Qy      378 ProThrGluLeuAenSerGluValLeuGlu-----GlnSerGln 390
Db      553 CCTGGAGACAGAGCTGCGCCAGCTGAGAGCTGGTGGGCGGAGCCAGAGCTTAC 612
Qy      391 GlnSerThr-SerLeuThrPhe-GlyGluGlyAaGluSerProGlyGly-----406
Db      613 GAGGCGCGGAGCGGACGTCGTCGAGCGGCTCATGACCTGAGAGAGATCAACAT 672
Qy      407 -----GlnSerCyAerArgGlyLysGlu-----414
Db      673 ACCATCTTACCGACGAAGAAAGATGCTGCTGTCGTCACAGCCATTTGGCTGTACATG 732
Qy      415 -----ProI 416
Db      733 CGCCACTTGGGCTGGCGAGCAAGAGTGAAGCAAGAGTGGGAGAACTTTCTCCG 792
Qy      416 yValSerIaIaGlyIaIaProGly-----SerHisProSerProIaIaProArgArgSer 433
Db      793 AAAAAGTGATGGGAAACCGCGGTGCGAGGACCTC-CCAAAGCAAGAGGCGTGAAG 851

```

```

QY 434 -----ProbaGlyAsnYsgIn-AlaAspGlyAlaAlaSerAlaPro 447
DB 852 CAGCATCTGGATCCGCCGCTGAAACCGGGAGAGCCCGAGTTCCAGATTTTCACA 911
QY 448 ThrGluGlu-GluGluGluValValYsAspArgThrGlnGlnSerYsProGlnProP 467
DB 912 CCTCAAGACAGAGGTGATGCCAGAGAGCCAGGTGTCACAGACGGAAAGGAGCGGG 971
QY 467 ogInProSerThrGlySerLeuThrGlySerPro-----HisLe 480
DB 972 GATGCCCTCTCGGAGCCGGAATATCGGGGCTCTGGCGAGACACCCCTGAGTCTCT 1031
QY 480 userGluLeuSerValAsnSerGln----- 488
DB 1032 GCACCTCTGTCTCTGGACACCCAGACCGGGAACCAAGCTCTACGCCCTCGAGCT 1091
QY 489 -GlyGlyValAlaProAlaAsnValThrLeuSerProAsnLeuSerPro-----AspTh 506
DB 1092 GGGGGACTCATCCCGGAGGGCCCAATGAGCTGAGATCCTTGCGGCCCGCAGAGATAC 1151
QY 506 rLySgAlaAspProLeuLeuSerProLeuLeuAsnArgGlnAlaCySProArgThrAs 526
DB 1152 CGACGAGGGGCGCAAAACGAGAGCCCGAGCTGAGATGAGGGGAGCGGGCGGCTC 1211
QY 526 p-----AspGluAspGluGlyArgArgGlyAspArgPheProThrArgLy 540
DB 1212 GGGAGCTGAGCTTAAACCAAGAGAGCTCCCGGCTGGCGAGACTCTCCCGCAGACAC 1271
QY 540 AlaArgPheLeuAlaYs-----GluValSerTh 550
DB 1272 CCTGCACAGCTGCGCCAAAGCCAGGTGAGCGGAGAGGATCATCAGCAGCTGCTGGT 1331
QY 550 rThrGluArgThrYrLeuYsAspLeuGluValIleTherSerThrPheGlnSerThrVa 570
DB 1332 GACAGAGCGCGCCACGCTGCATCGCGGTCTGCACACCTCTTCTCCAGCCCAT 1391
QY 570 lSerLyGluAspAlaMetPro--GluAlaLeuYsSerLeuLeuPheProAsnPheG 589
DB 1392 GGCAGAAATGCTGTCTTCCCTTGGAGAGACTCAGAAC--ATCTTCCCGACCTGGA 1448
QY 589 uProLeuNlYsYrPheNlYrThAsPheLeuYsGlnlGluGlnArgLeuAlaLeuTr 609
DB 1449 CGACCTCATGAGGTGATCCCTGCTCTGATCGGCTGTGAAGCGG----- 1497
QY 609 pGluGlyArgSerAsnAlaGlnIleArgArgTr-----GlnArglGluYsArgValMe 627
DB 1498 -----AGGCAAGAGAGTGGCTATCTATCGAGAGATCGGAGACGTGCT 1541
QY 627 tLeuYsAsnIleGlnGluMetYsNlYsLeuAlaAlaHisYsLeuTr----- 642
DB 1542 GCTGCGCGCGGTTTATGCT-----GCTGAGGCGCTCGTGTCCAGAAATCTC 1589
QY 643 -----LySHisSerGluAlaLeuGluAlaLeuGlnYsGlnYsLeYsSe 657
DB 1590 CTCCTCGCTTGCAGCCGCGAGTATGCTTGTAGACAGCTCAAGCAAGCAAGCAAG 1649
QY 657 rSerArgArgLeuGluAsnPheCySArgArgPheGluLeuGlnYsValYs--TyrLe 676
DB 1650 GGAACCTCGGTTCTGTGCTGCTGCGAGAGCTGAGAGCGCGCGGCGGCGCGCT 1709
QY 676 uProLeuAsnThrPheLeuLeuArgProLeuNlYsArgLeuMetNlYrYrLySgValLe 696
DB 1710 GCAGCTAAAGACATGATCCCAAGAGATGACAGGCTGACCAAGTACCCCGCTCTCT 1769
QY 696 uGluArgLeuCySValYsNlYs--HisProSerSerHisAlaAspPheArgPhe 712
DB 1770 GCAGAGATCGGCGGAGAAACAGAAAGCCCAAGAAACGGAATGAGCTGCGAGC 1829
QY 713 -----CyArgGluAlaLeuAlaGlnIleThrGluMetValAlaGlnLeuNlYsGlyTh 730
DB 1830 CGAGTGTGCTGGGAAATTTTACACACAGTCAACCAAGCGCTGCTGATCATGAG--GA 1886

```

```

QY 730 rMetIleYsMetGluAsnPheGlnYs-----LeuNlGluLeuYsYsAsP-- 746
DB 1887 CTGTCTAGGCTCAAGACATATACCGCGGCTCGACTGTCTCCACCTTGCGCAGACAG 1946
QY 747 -----LeuNlGlyIleAspAsnLeuValValProGlyArgGluPheIleArgLe 763
DB 1947 GCACCTATGCTAGAGAGATTCAGAAACCTGTGACATCCCAAGAAATTTGTTCCACA 2006
QY 763 uGlySerLeuSer--LySLeuSerGlyYsGlyLeuGlnGlnArgMetPhePheLeuPh 782
DB 2007 GGGCCCACTGACTGCGGGGTGACTAAGCAAGACAGTGAAGTGCATGTGCTGCTGCT 2066
QY 782 eAsnAspValLeuLeuYrThr-----Se 790
DB 2067 GAGACACTGTCTGTCTGTCTGTCAGAGCCAGACAGACGCGGCTGTCTCAATGCCATAG 2126
QY 790 rArgGlyLeuThrAlaSerAsnGlnPheYsValHisGlyGluLeuProLeu----- 807
DB 2127 CCGGACACTGACCCCGCAGCCCGATGCAAGACCATGTCTGCGCCGTGCGGCTCAC 2186
QY 808 -TyrGlyMetThrIleGluGluSerGluAspGluTrGlyValProHisCySLeuThrLe 827
DB 2187 CTCGCCCATACCCGAGGCTGCGCACCCATTCACAAAGCTTCTAGCTCTTTTACTCTG 2246
QY 827 uArgGlyGlnArgGlnSerIleIleValAlaAlaSerSerArgSerGluMetGluYsTr 847
DB 2247 GAGACAGAGAGCCCAATATACAGCTGTGCGCACAGACTGTGCGAGCGGAAATCTG 2306
QY 847 p-----ValGluAspIleGlnMetAlaIleAspLeuAlaGluYsSerSerPr 864
DB 2307 GTGTGCTCTCATCATGAGACTGCGGATCCCGAAAGTCCCTGCCCTCTGCGCC 2366
QY 864 oAlaPro-----GluPheLeuAlaSerProProAspAs 876
DB 2367 TAAGCCCGCGCCAGCGCCAGACAGACCCGAGAACCTTCTTACAGACTCTAGAACG 2426
QY 876 nLySerProAspGluAlaThrAlaAlaAspGlnGluSerGluAspAsPheSer 894
DB 2427 GAATGTGCGCGAGAGACGTCCTCAGCTGATGCCGAGACCGAGAAATCTTCACT 2481

RESULT 5
US-60-796-903-635
; Sequence 635, Application US/60796903
; GENERAL INFORMATION:
; APPLICANT: Johnson, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Colorectal Cancer Microarrays and Methods of Use
; FILE REFERENCE: 55815-300 (327365)
; CURRENT APPLICATION NUMBER: US/60796,903
; CURRENT FILING DATE: 2006-05-02
; NUMBER OF SEQ ID NOS: 1071
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 635
; LENGTH: 4768
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-60-796-903-635

Alignment Scores:
Pred. No.: 0.0226 Length: 4768
Score: 174.00 Matches: 120
Percent Similarity: 36.9% Conservative: 96
Best Local Similarity: 20.5% Mismatches: 229
Query Match: 3.2% Indels: 143
DB: 8 Gaps: 27

US-09-555-342B-2 (1-1045) x US-60-796-903-635 (1-4768)
QY 365 HisSerLySerLeuHisSerIleArgSerLeuAlaSerGluProThrGluLeuAsnSerGlu 384
DB 47 CATGAGACCTGGAACAGATGCGCGAGTGGCT---CATCTATTGCA----- 91

```

```

Oy 385 ValLeuGIInSerGIInSerThSerLeuThrPheGlyGIInGIValAGIuSerPro 404
Db 92 -----GGTGGTGGCCACCAACCG 112
Oy 405 GlyGIInSerCyArGArgGlyLyGIuProLyValSerAGIuGIuProGIuSer 424
Db 113 GGTACCTGGGACCTCGCTCAGGTGTTCAGCT-----TGGCGAGACCTCCGCGAGTGG 166
Oy 425 HisProSer---ProAlaProArGArgSerProAlaGlyAenLyGIInAlaSerGly--- 442
Db 167 AGTCTGCTCGCCAGCTCTTAACAACTCCGGCGACCTCCATCAACTGAAGAGAT 226
Oy 443 -----AlaAlaSerAlaProThr-----GIuGIuGIInGIuGIValLy 456
Db 227 CAACCTAGGCCCGCAGATGCTCCAGTTCTCTGTGGAAGAACTAGAGCACTTTC--- 282
Oy 457 AspArgThrGIInSerLyGIuProGIInProGIInProSerThGIuSerLeuThrGI 476
Db 283 -----TCACGGCTGTGTGAGACGTTGGAAATGAGAAAGTAACTTT 327
Oy 477 SerProHisLeuSerGIuLeuSerValAsnSerGIInGIVal----- 491
Db 328 TCGAGGCACTTTGACTTG-TTTGATGTTGTGACTTTGGAAAGTTATAGAAACATTATCA 386
Oy 492 -----AlaProAlaAsnValThrLeuSerProAsnLeuSerPro----- 504
Db 387 CGACTTCTCGAACACCTATAGCATTTGGCCACAGGAATCAGGCCCTTCCCAACAGAA 446
Oy 505 -----AspThrLyGIInAlaSerPro-----LeuLeuSerProLeu 517
Db 447 AGCATTAATGATGAGACATCTACAAAGCGCTTCTGATTAAATGATGAAACCTTGTG 506
Oy 518 AsnAspGIInAla-----CyProArGThrAspArgGly----- 528
Db 507 GAAATGAAAGAAATCTCTATGACCTGTGTTATGGGAGATGAAAGTGAAGAACTTAT 566
Oy 529 -----AspGIuGIuLyArGArgLyAspArgPro---ThrAspLyAla 541
Db 567 GAGACCTTAATGAAGGACAGAGAGACACATCAGCCCAATGTCCGAAATGATTTACGA 626
Oy 542 TyrPheIleAlaLyGIuValSerThrThrGIuArgThrTyLeuLyAspLeuGIuVal 561
Db 627 AGTTGTGTCTAGCAGAAATTAAGACAGACAGAAATTAATACAGAAACTTGGAGTCA 686
Oy 562 IleThrSerThrPheGIInSerThrValSerLyGIuAspAlaMetProGIuAlaLeuLy 581
Db 687 ATAGAAAAGTATTTTCATGGCACCACTAAAGAA-----TTTCTGACAGCAGCAGAAATTT 740
Oy 582 SerLeuIlePheProAsnPheGIuProLeuHisLyPheHisThrAspPheLeuLyGIu 601
Db 741 GATTCACTATTCAACATTCCTGCACTGTAAACTTCAATCGGAACCTTAATGCAGAG 800
Oy 602 IleGIuGIInArgLeuAlaLeuTrpGIuGIuLyArgSerAlaGlnIleArgAspTyGIIn 621
Db 801 ATTCATGAT----- 809
Oy 622 ArgIleGIuAspValMetLeuLyAsnIleGlnLyMetLyHisLeu----- 637
Db 810 -----TCCATTGTAAATATAAATAAGCAAGAACTGTACCAAGTTTATTATTAACATAC 860
Oy 638 AlaAlaHisLeuTrpLyHisSerGIuAlaLeuGIuAlaLeuGIuAsnGIuIleLeuSer 657
Db 861 AAGGAAAGATTGGTTATTACGGGCAATACGACGTGAGAGAGATCGACCACTCTTACT 920
Oy 658 SerArgArgLeuGIuAsnPheCyArGArgPheGIuLeuGIIn---LyValCyArTyGIu 676
Db 921 TTAGACTACATTTCTAAGTCAAAAGAGATGTCAAACTGAATATAGAGAAATGTTCCAA 980
Oy 677 ProLeuAsnThr-----PheLeuLeuArg-----ProLeuHisArgLeu 689
Db 981 AGACCAAAATTAATGGAAATTAATCTTGAAGACTTGTGCTGTCCATATCAACGTGTT 1040
Oy 690 MetHisTyLyGIInValLeuGIuArgLeuCyLyHis-----HisProProSerHis 707

```

```

Db 1041 TTAAGTACCACTTCTCCTCAGAGAACTGTCAAACATACCACTGATCCGATGAGAG 1100
Oy 708 AlaAspPheArgAspCyArGArgAlaAlaLeuAlaGIuIleThrGIuMetValAlaGlnLeu 727
Db 1101 GCAATCTGAAACCTGGCTCTTGATGCCATGAAAGACCTTGCAATATGTAAATGAA--- 1157
Oy 728 HisGIuThrMetIleLyMetGIuAsnPheGIInLySerHisGIuLeuLyAspLeu 747
Db 1158 -----GTAAGAAGATTAAT---GAGACCTTCTGTAAATTAACAGTTTCAG 1202
Oy 748 IleGIuIleAspAsnLeuValPro-----GlyArgGIu 759
Db 1203 CTACTATATAGGAATTTGAACCAACAGTTTGTCTTTTGGACGACCTCAGGAGATGCT 1262
Oy 760 PheIleArgLeuGIuSerLeuSerLyGIuSerGIuLyGIuLeuGIInArgMetPhe 779
Db 1263 GAAATTCGAATTAACACTCTGACAGCATAC-----AAACAAAGAAAGCATATTC 1313
Oy 780 PheLeuPheAsnAspValLeu-----TyrThrSerArg 791
Db 1314 TTTTATTTGATTTGGCAGTATGCTATGTAAGAAAGGTGATTAATGAAATGAAG 1373
Oy 792 GlyLeuThrAlaSerAsnGlnPheLyGIuValHisGIuGIInLeuProLeuTyGIuMetThr 811
Db 1374 GAAATTAATAGTCTTTCAGCAGTACAGATACCAAT-----AATCTTACA 1418
Oy 812 IleGIuGIuSerGIuAspGIuTrpGIuValProHisCyLyLeuThrLeuArgGlyGlnArg 831
Db 1419 ACCCATTAAGAAACAAAAAGTGTCTTATGCTTACCTTACCTCAACCAATACCAAGACAA 1478
Oy 832 GlnSerIleIleValAlaAlaSerSerArgSerGIuMetGIuLySTPValGIuAspIle 851
Db 1479 AATGGTTAGAAATTTATTTGCAAAACAAAGATTTAAAGAAAGATGCTAGAACAGTTT 1538
Oy 852 GlnMetAlaIleAspLeuAlaGIuLySerSerSerProAlaProGIuPheLeuAlaSer 871
Db 1539 GAATGGCTTG-----TCTAACATTAACCAACGACTATGACACTCC 1580
Oy 872 SerProProAspAsnLy 877
Db 1581 AATTTCCACGACTTCAAG 1598

RESULT 6
US-60-796-903-638
; Sequence 638, Application US/60796903
; GENERAL INFORMATION:
; APPLICANT: Johnston, Patrick
; APPLICANT: Harkins, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Colorectal Cancer Microarrays and Methods of Use
; FILE REFERENCE: 55815-300 (327365)
; CURRENT APPLICATION NUMBER: US/60/796, 903
; CURRENT FILING DATE: 2006-05-02
; NUMBER OF SEQ ID NOS: 1071
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 638
; LENGTH: 13664
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-60-796-903-638

Alignment Scores:
Pred. No.: 1.29 Length: 13664
Score: 154.00 Matches: 205
Percent Similarity: 33.1% Conservative: 137
Best Local Similarity: 19.9% Mismatches: 347
Query Match: 2.8% Indels: 345
DB: Gaps: 48

US-09-555-342b-2 (1-1045) x US-60-796-903-638 (1-13664)
Oy 89 LysLysIleThrValTrpLeuAspLeuLeuAspPro-----IleValLyGIInIleArg 106

```

Db	1363	AGAAACCTCAGAGTTTAT-----CTGAAACCTTGAAATTAATCTATTTCAGAAAGG	1433
QY	107	ArgProLYSHLeVal-----ValValLYPheValValLYSPhe	119
Db	1434	AGATCAAAAGAGAGTCAAGAAAGAAAGAAAGACACTTTTAAAAAGATGCCAACAGTA	1493
QY	120	PheProPheSPHLeThr-----GlnLeuGln	128
Db	1494	TTCTCTCCAGCAAGGACCATTAAGAGAGTGTGAGTTGTGGATGACATAACTACTAGA	1553
QY	129	GluGluLeuThrArgTYrLeuPheAlaLeuGlnValLYSGlnAspLeuAlaGlnLYArg	148
Db	1554	GAGAGCCTT-----ATGATTCAGTTGAAGAA-----	1580
QY	149	LeuThrCYaAsnAspThrSerAlaAlaLeuLeuLeuSerHisAlaValGlnSerGluLe	168
Db	1561	-----GTACTGAAAGTGAATCACTCACTCTTCCACAAAGAA-----	1616
QY	169	GlyAspPheAspGluAlaLeuAspArgGlnHisLeuAlaLYaAsnLYeTYrLe-----	186
Db	1617	-----AACGAAGAACTAGTGTGGTCTATCAAGAGr-GATGAGTATTAACAGAA	1666
QY	187	-----ProGlnGlnAspAlaLeuGlnAspLYeAlaGlnPheHis	200
Db	1667	GACAGCCAAAGATCAGGGGAAGAGAGACCTTCATTCAAAACAATCATGAATGAGT	1726
QY	201	HisAsnHisLeGlyGlnThrProAlaGlnSerAspPhe-----GlnLeuLeuGln	217
Db	1727	GCTGTGAAGATGCTTCACAGCTTCGAGATGATGATGATGATTCATCAGAAACCACTCAAGAG	1786
QY	218	IleAlaArgArgLeuGluMetTYrGlyLeaArgLeuHisProAlaLYaAspArgGluLY	237
Db	1787	ATTGTACACAGTACTGAG-----AGCCAGAAAGAGAGAA	1822
QY	238	ThrLYeIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLYe	257
Db	1823	ACAAAG-----GGAAAAAGCCAAA	1840
QY	258	IleAsnAlaPheAsnThrAlaLYeValaArgLYeSerPheLYeArgLYe-ArgPheLe	277
Db	1841	ATTAACAGCA-----GGGAACACACAGCACTGACGAAAAAGTACACAAAGCATT	1891
QY	277	uIleLYeLeuArgProAspAlaAsnSerAlaTYrGlnAspThrLeuGlnPheLeuMetAl	297
Db	1892	GATGAAGATGCAGAAAGACGTCACTCAGCGCATGATGAAGACGATGAAGCATTTGATGA	1951
QY	297	aSerArgAspPheCYeLYeSerPheTYrLYeCYeValaGlnHisAlaPheAsr	317
Db	1952	AGTCTCAACTT-----	1963
QY	317	GluPheGluGlnUProLYeProLYeProLYeProValLeuPheSerArgLYeSerSerPh	337
Db	1964	AAATATCAGAAACTTAATAAGTCAGAAAGTGAAGAACTTGTACTGAGAGAA-----	2017
QY	337	eArgPheSerGlyArgThrGlnLYeGlnValLeuAspTYrValLYeGluGlyGlyHisLYe	357
Db	2018	-----GAGGGCTACG	2028
QY	357	AlaValaGlnPheGluArgLYeHisSerLYeIleHisSerIleArgSerIleuAlaSerG	377
Db	2029	CGGATTT-----AA	2037
QY	377	nProThrGluLeuAsnSerGluVal-----LeuGluGlnSerGlnGlnSe	392
Db	2038	AACAATTAGCTCAACAGTACATATGACATATAATATTCTCAGAGTCATCACAGAAAA	2097
QY	392	rThSerLeuThrPhe-----	397
Db	2098	AACAAGTTTGTATTTTTGACGAAGACCAGATTCGAATTCGAAGACCTCAGACACTACC	2157
QY	398	-----GlyGluGlyAlaGlnSerProGlyGlyGlnSerCYeArgArgG	412

[illegible]

```

Oy      728  eglYthrtMetIle---LysMetGluAsnPhelInlyLeu----- 740
Dy      3170  -GGTACTATGGTAAGTCTTCTACAGAAAGAAAGAAAGAACTAGATGCTGATGCTGCTTA 3228
Oy      741  -Hl eglYleuLybLYbAspLeu----- 747
Dy      3229  TGAAGAACTTTTATGAAAGAGCAACAAGATGCATTAACCTGGATGCTTAACCCAAACCCAGGC 3288
Oy      748  ----l eglYl leaAspAnLeuValValProglYArgGluPheIl eArgLeu----G1 764
Dy      3289  CCCCATTGGT---GAGGATATGACAGAGTCCACCATGACATTGACAAATGCCAGATGCG 3345
Oy      764  ySerLeu-----SerLYbLeuSerG1LYb eglYleuGlnGlnArgMetPhePheLeuPh 782
Dy      3346  CTCTTTGACATCAAMGTGTTCTCTCAGAGACGCTCTTTCACAGATTGC----- 3391
Oy      782  eAsnAbpValleuLeuYrthrSerArgG1YleuThrAlaSerAenGlnPheLYbValH1 802
Dy      3392  -----ACCAAGAGTGAACAACACTCTATGCC 3417
Oy      802  eglYGlInLeuProLeuYrG1YMetThrIl eglGluSerGluAbpGluTrrG1YValPr 822
Dy      3418  AGATGTTAAATTAACCCAAACATTTTTCACACAGAAATAATGAGATGATATGTAAACCGA 3477
Oy      822  oHl eCybLeuThrLeuArgG1Y-----GlnArgInSerIleIleValAlaIaIaSerSe 840
Dy      3478  TTATACAAAGAGAAATTCACAGATATTAATGCCCATGATAGCGGATTTTGACCTACTCGGA 3537
Oy      840  rArgSerGluMetG1uLYbTrrValGluAspIl eglInMetAlaIleAspLeuAlaGluLY 860
Dy      3538  GCCTTCAGAAAGTGCTACATCTGTC-----CCACC 3567
Oy      860  sSerSerSerProAlaProGluPheLeuAlaSerSerProProAspAnLYbSerProAs 880
Dy      3568  CTCGACACACACCTTCTCACAATCATCTGTTCTTCGGTCTGTACCAAGATAGCTCTTC 3627
Oy      880  pGluAlaThrAlaAlaAspGln----- 887
Dy      3628  ACCCATTTACTACCTCGGATAGCATATACACAGCTTATATACAGACCAAGTGCACATGATAC 3687
Oy      888  -----GluSerGluAspLeuSerAlaSer----- 896
Dy      3688  TAAATTTGAAGATTCTGAGAAATTTCTTCATCAACTTATTTCCAGCAGACATTATAGA 3747
Oy      897  -----ArgThrSerLeuGluArgGlnAlaProHl bArgG1YbAsnThrMe 911
Dy      3748  CTATCCAGAAAGAAATTAAGTGTATCTTTTGATGTCGACCTCCCAACAGATGTAGAGCTAG 3807
Oy      911  tVal---Hl eValCYbTrrPHl bArgAsnThrSerValSerMetValAspPheSerIleAl 930
Dy      3808  TGTGATCATATTTGTTATTTCTTATTCCTTATTCGATATAGGCATCTTCTATCATAGAAATCTG 3867
Oy      930  a-----ValGluAsnGlnLeuSerG1YbAsnLeuLeu----- 940
Dy      3868  ACCTAAACCTGAAGGCCACGATTGCTGCACAGCTGTTCTATCTGACTTACTTATATCTGAAA 3927
Oy      941  -----ArgLYbPheLYbAsnSerAsnG1Y 948
Dy      3928  GGACCCAGTGAAGAAAGCCAAAGAGAAACTGGG 3961

RESULT 7
US-10-529-655-22
; Sequence 22, Application US/10529655
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research, et al.
; TITLE OF INVENTION: Human Sarcoma-Associated Antigens
; FILE REFERENCE: 10461, 70156M00
; CURRENT APPLICATION NUMBER: US/10/529, 655
; CURRENT FILING DATE: 2005-03-30
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22

```

```

:      LENGTH: 5676
:      TYPE: DNA
:      ORGANISM: homo sapiens
US-10-529-655-22

Alignment Scores:
Pred. No.:      3.39      Length:      5676
Score:      139.50      Matches:      205
Percent Similarity:      34.0%      Conservative:      146
Best Local Similarity:      19.08      Mismatches:      425
Query Match:      2.6%      Indels:      258
DB:      6      Gaps:      43

US-09-555-342B-2 (1-1045) x US-10-529-655-22 (1-5676)

QY      2      G L Y G I I I E G L I G I N A T G P R O T H R P R O G I Y S E R A R G L E G I A L A P R O G I L U A S N S E R G L Y      21
Db      1345      G A G A A A A C A T G A G A A A G A G A C G C T G T T C T C T G T G A A T C T T G A A T G C A A G C G C      1401
QY      22      I L E S E R T H R L E A G I A R G G L Y G I L Y A P R O P R O T H R P R O S E R G I L Y L E U V A L S E R      41
Db      1402      G T C C A G G T A C C C C C C A C A C A G A A A T A G A G A C T G A G A G A A G T G C A A A G A A C T T      1466
QY      42      I L E Y S I L E G I M E T L E U A S P A S P T H R G I N G L A L A P H E G I U A L P R O G I N A R G L A P R O      61
Db      1462      G G A A A G T G T A C T G T T G T T C A G A T G A T A G C      1506
QY      62      G L Y L Y S V A L L E U L E U A S P R A L A V A L C Y S A S N I S      76
Db      1507      T E T A T A T C A C A A G C T G A A T T A G C T T T A G C A T G A T T      1542
QY      77      G L U G L Y      A A P T Y R P H E G L Y L E U G L I U A P H E P R O A S P H I S L Y S I L E T H R V A L T P R L E U      95
Db      1543      A A A G G T C T T G A T T A A G C C C A G C C A G C C C C G A G G A C A G C G T T C G A T T C A G T G      1602
QY      96      A S P R I E U E U S Y P R O L I E V A L L Y S G I N      104
Db      1603      A A A T C C A G C C T A C C T G A G C C A A C C T G G C C C T A G C A T G A C A G A T G A G T T A G T C C G G T      1667
QY      105      I L E A R G A T P R O L Y S H I E V A I V A L I Y A L L Y S P H E V A L I V A L L Y S P H E P R O P R O A S P      123
Db      1663      T T C C T T A A C A G G C T T T G A A A C C C T T T A C A A C A C C T G I G A G      1719
QY      124      H I S T H R G I N L E U G I N G L Y      G L U L E U T H R A R G Y L R L E U P H E A L L E U G I N V A L      140
Db      1720      C T T T C A G C C T G C A G A G C T G T G G A G A T C T C T G T G A A G A T T A T T T G C C C C T C G G G T C      1779
QY      141      L V S G I N A S P R L E U A L S G I N I L Y A R G L E U T H R C Y A S N A S P T H S E R A L A      156
Db      1780      C A G C C C A T A G T A A A G A G T C C T A A A C A A A A A T C C A G A A C A G G C C A C A A C A G A T T      1833
QY      157      A L A L E U L E U I L E S E R N I S I L E V A L G I N S E R G I U I L E G I A S P R P H E A S P G I U A L E U A S P      176
Db      1837      A T A A C T C A G A G C T G T A T C T G A T      1860
QY      177      A R G G I N H I S L E U A I A L Y S A S N L Y E T Y R I L E P R O G I N G I N A S P R A L A L E U G I A S P L Y I L E      196
Db      1861      T C C C A C T T G G A G A A C T C C A G A A A A A A A T C C A G A A C A G A G C C A C C A A C A G A T T      1917
QY      197      V A L G I U P H E N I S H I S A S N I S I L E G L Y      205
Db      1918      C T T C A A G A A A C T T A T A T G A A T A G A C T A T A A C T T A A A G T G C T C A G A G A T C G T C T C A A      1977
QY      206      G I N T H P R O A L A      209
Db      1978      A A G C A A G A T G T A C A A T T C A G A A C C T A A G G A A A C T G A A A A C A G G A A C T G A G A C T      2037
QY      210      G L U S E R A S P R P H E G I N L E U L E U G L Y      I L E A I A R G A R G L E U G I U M E T      224
Db      2038      G A G A A G T T T A C C A G A C T T T G A A A G C T C A A A A T G A C A C A A T G C A A A G C C T T C G A A A T G      2097
QY      225      T Y R G I L I L E A R G L E U H I S P R O A L A L Y S      A S P A R G L E G L Y L I T H R L Y S      239

```


Db 2098 -----CTGCACCAAAAGCCAGCTTGGACAACTTCACAGCTCAGAGGACT--- 2142
 Qy 240 ILeuLeuAlaValAlaAenThrGlyIleLeuValPheGlnGlyPheThrIleAsn 259
 Db 2143 ---TCTCAGCTTCAGCAACAGTACGTCTGCTTATCTTCAGAGTCTTATTTTCGACG 2199
 Qy 260 AlaPheAnThrAlaValArgValArgLeuSerPheIleValArgPheLeuIleLys 279
 Db 2200 CAACCTGAATATACAGAAAGCTCCAGAGGTTGTACGACAGAAAGAGCCCAACTG----- 2253
 Qy 280 LeuArgProAspAlaAenSerAlaTyrGlnAspThrLeuGlnPheLeuMetAlaSerArg 299
 Db 2254 -----CCTGATGCAAAACAATGTGTGCAATTTGTAGAGCTGACGACACAGAGAGTGA 2307
 Qy 300 AspPheCysIleSerPheThrIleValIleCysValGlu---HisIleAlaPhePheArgLeu 318
 Db 2308 CAGACGAAAGAGCTTCTTGGAAACATACAGAGAAATTCGCAAAAGCTTCGACAGACTA 2367
 Qy 319 PheGlnIleProIleProIleProIleProIleProIlePheSerArgIleSerPheArg 338
 Db 2368 CAAGAAAGAAATTCGAAATTAAGAGCCAAAGCTTCGTGCTGGAGGCTGAAAAATACAT 2427
 Qy 339 PheSerIleArgThrGlnIleValLeuAspTyrValIleValGlnGlyIleValLys 358
 Db 2428 ---GAGATTGAAAGCCAGAAACAACATCCAGACCTAAACCATGATCTGAGTCACAG 2484
 Qy 359 ValGlnPheGlnArgIleValSerIleValIleSerIleArgSerLeuAlaSerGlnPro 378
 Db 2485 GAGCAGTTGCTTCAGAAATTCGGAGCTCTTACAGATGCAATGTAATCTACAGCAAAAC 2544
 Qy 379 ThrGlnLeuAsnSerGlnValLeuGlnIleSerGlnIleSer 392
 Db 2545 CTTGAAACAAATGAATGTTGCTTGAGAAACTTCGCCAGCAATACATGATTAAGCTGTT 2604
 Qy 393 -----TherLeuThrPheGlnGlyIleGlu 402
 Db 2605 GCTCTGGAGCGGCTATATAGTAAATAATCTGCTGCTAGAGAAAGAAAGAA-AGAACT 2663
 Qy 403 SerProGlnIleGlnSerCysArgArgIleGlu----- 414
 Db 2664 GCCGCACTTCCTTCTGCTGTGAGAGCCAGATCAGACTTACAGAGCTGCCGATGT 2723
 Qy 415 ProIleVal-----SerAlaGlnProGlnSerHisProSerPro 428
 Db 2724 CCTCTCTCAATGAAGCTACTATGCAAGATGAGAG----- 2762
 Qy 429 AlaProArgArgSerProAlaGlnValAsnIleGlnAlaSerGlnAlaIle----- 444
 Db 2763 -----TCTCTGAGAGGCCAAAGGCTGGAAGTGAACAGTTATCTACTACTG 2810
 Qy 445 SerAlaProThrGlnGlnIleGlu-GlnGlnValValIleAspArgThrGlnIleSerIlePyr 464
 Db 2811 TCAAAACCTCCAGTGGGTGAAGAGAAATG-----GAAACCAATTTTGGCCGTTG 2861
 Qy 464 GlnIleProGlnIleProSer---ThrGlySerLeuThrGlySerProHisLeuSerGlu 483
 Db 2862 GCAGAAAGAAACAAGAGATCATTCAGCAGTTACAGACGCTCTCTTAT----- 2910
 Qy 483 userValAsnSerGlnIleGlnValAlaProAlaAsnValThrLeuSerProAsnLeu 503
 Db 2911 -----GATAGGAACAAGAGTGAAGATCTTAAAGCAACATGCTTCGCAAACTTGG 2963
 Qy 503 rProAspThrIleGlnIleAlaSerProLeuIleSerProLeuLeuAsnAspGlnAlaCysPyr 523
 Db 2964 ACCAGGCGCAGAGTGA-----ATAGCAAGAGAGCTGTCCA 2999
 Qy 523 oArgThrAspAspGlnAspGlnIleArgValArgValArgPheProThrAspIleValTyrPh 543
 Db 3000 GCGCTCTACAG-----CGAAAGAAAGATGCTGCAGAGACTT----- 3036
 Qy 543 eIleAlaIleGlnValSerThrThrGlnArgThrThrIleLeuIleAspLeuGlnValIleTh 563
 Db 3037 -----CTAAGTATCGAAATTAACAAGTCTGTGAACATGAAATGAGATTCA 3083

Qy 563 rSerTrpPheGlnSerThrValSerIleGluAspAlaMetProGlnValLeuIleYserIle 583
 Db 3084 AGGCGCTCTTCACTCTGTGTAGACACAGGAGCAGAAAGCAAGCTCTGCAGAGAAATT 3143
 Qy 583 uIlePheProAsnPheGlnProLeuHisIleValPheHisIleThrAsnPheLeuIleGln 603
 Db 3144 GGTGCAAGCTTATATGAA-----AGAAATTCAGAAATTAACAGGCTCTGCG 3188
 Qy 603 uGlnIleGluAlaLeuTrpGlnIleArg-----SerAsnAlaGlnIleArgAs 619
 Db 3189 CCAATATTATGGA-----GGAGAGACTCCCTGATGTCCCAAGCACCTTCTTAA 3239
 Qy 619 pTyrGln-----ArgIleGlyAspValMetLeuIleValIleGln 632
 Db 3240 CCAACAAAGCTGAAGTTACCCCACTGGCGCTTTGGAAG-----CAGACTGATCA 3290
 Qy 632 nGlyMetIleValLeuAlaHisIleuTrpIleValSerGlnAlaLeuGlnAlaLeuGln 652
 Db 3291 AGGTTCAATGCAGATACCTTCCAGA-----GATGATAGCACTTCAATGCTGCAGAA 3344
 Qy 652 uAenGlyIleIleYserSerArgArgIleGlnIleAsnPheCysArgAspPheGlnIleGln 672
 Db 3345 GGAATGCAGATCCCAAGATCCACATTAAGGACTTGGACACAGCTTCAGAGCTGCAAAA 3404
 Qy 672 sValCysTyrLeuProLeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisTyr 692
 Db 3405 A-----GAACTGAGTATATGC 3419
 Qy 692 rIleValIleGlnIleGlnIleCysIleValHisIleProProSerHisIleAlaAspPheArg 712
 Db 3420 CAAAGGAACTTGAACCTCATGCTGCTAAAGAA-----AGAGA 3458
 Qy 712 pCysArgAlaAlaLeuAlaGlnIleThrGlnMetValAla-----GlnLeuHis 728
 Db 3459 AAGTCAGATGAAACTTTCCTCTTACAGTCCATGATGGCTGTGCAGAGAAAGAGCTGCA 3518
 Qy 728 sGlyThrMetIleIleYMetGlnAsnPheGlnIleValLeuHisGlnLeuIleValAspLeu 748
 Db 3519 GGTGCAAGGCTGTGATATGAGATGCTGTGACACAGAAACATACAGATTAAAGAAATCTCAT 3578
 Qy 748 eGly-----IleAsnLeuValValProGlnArgIleGlnIleArg 762
 Db 3579 AAAGACCTGCAATCAACACTGTGATCTCTGAAGACATACAGCATATGGA----- 3630
 Qy 762 gLeuGlySerLeuSerIleLeuSerGlyIleGlnIleArgMetPhePheLeuPh 782
 Db 3631 -----GCCCTGACCCAGAAAGTCTTACTTCTTCG 3659
 Qy 782 eAsnAspValLeuLeuTyrThrSerArgIleLeuThrAlaSerAsnGlnPheIleValHis 802
 Db 3660 GGAAGAAAGTTGCTTCAGTGAATCCAGAGGTCAAGAAATTTACAGAAACCGAAGACAA 3719
 Qy 802 sGlnIleLeuProLeuTyrGlyMetThrIleGlnIleGlnIleAspGlnIleTrpGlnValPyr 822
 Db 3720 GTTCTGCTGTATCTGAAGAGCTAGTATGAAGCAAGAGTGG-----CTCAA 3767
 Qy 822 oHisCysLeuThrLeuArgGlyIleArgGlnSerIleIleValAlaAlaSerSerArgSe 842
 Db 3768 TGAAGCTTACAAAGCAGAGACAGCTCTATAGCACTGTGTGAAGTTCATGCCATGCC 3827
 Qy 842 rGlnMetGlnIleTyrValIleGlnAspIleGlnIleMetAlaIleAspLeuAlaGlnIleYserSe 862
 Db 3828 AGAGACTCTGAGAGAGACCAAGACTCTGCAGGTGAAGTGAAGGGCTCAGGTGTACG 3887
 Qy 862 rSerProAlaProGlnPheLeuAlaSerProPro-----AspAsnIle 877
 Db 3888 CAGTTCGCTGAAGAAAGTTCTTGAAGAAAGCTTGAAGCCTTAAACAGCTGAGACCTT 3947
 Qy 877 sSerProAspGlnAlaThrAlaIleAspGlnIleGlnIleAspAspLeuSerAlaSerArg 897
 Db 3948 GGCCTGCAATTTGAGGTGCAGCTGCAGGGATGACACCGAAGATACAGCACTGATTCAC 4007

Qy 897 gThrSerLeuGluArgGlnAlaProHisArgGlyAsn 909
Db 4008 TGACAGTATTGAGAGAGAGCTGCACACCATATGTCAC 4044

RESULT 8
US-10-576-496-19
Sequence 19, Application US/10576496
GENERAL INFORMATION:
APPLICANT: Teijin Pharma Limited
APPLICANT: Komori, Toshihisa
APPLICANT: Kanatani, Naoko
APPLICANT: Yoshida, Caroline Andrea
APPLICANT: Zama, Akira
APPLICANT: Kobayashi, Shinji
APPLICANT: Yamane, Kei
TITLE OF INVENTION: BONE AND/OR JOINT-DISEASE-ASSOCIATES GENES
FILE REFERENCE: Q94468
CURRENT APPLICATION NUMBER: US/10/576, 496
CURRENT FILING DATE: 2006-04-20
PRIOR APPLICATION NUMBER: PCT/JP2004/015879
PRIOR FILING DATE: 2004-10-20
PRIOR APPLICATION NUMBER: JP 2003-359172
PRIOR FILING DATE: 2003-10-20
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn version 3.1
SEQ ID NO 19
LENGTH: 3276
TYPE: DNA
ORGANISM: Mouse
FEATURE:
NAME/KEY: CDS
LOCATION: (114)..(2763)
US-10-576-496-19

Alignment Scores:
Pred. No.: 5.35 Length: 3276
Score: 131.50 Matches: 122
Percent Similarity: 34.7% Conservative: 92
Best Local Similarity: 19.8% Mismatches: 197
Query Match: 2.4% Indels: 206
DB: Gaps: 30

US-09-555-342b-2 (1-1045) x US-10-576-496-19 (1-3276)

Qy 453 GluValValIysAspArgThrGlnGlnSerIysProGlnProGlnProSer----- 470
Db 43 GAAAGTGTGAGGACACAGAGCGACCTGCATACCTGAGCTCCGAGAGAGAGCTGCAC 102
Qy 471 -----ThrGlySerLeuThr-----GlySer 477
Db 103 TCTGGCATCCAGTACGACAGAGACTAAGCATAAAGAGCCATGACTTCTGCGCCT 162
Qy 478 ProHisLeuSerGluLeuSerValAsnSerGlnGlyValAlaPro----- 493
Db 163 AGTCATATGCTGGAAGCTGCCCTGGAGACAGATGACGGGATCATTCAGAGCACTAAACA 222
Qy 494 -----AlaAsnValThrLeuSerProAsnLeuSerProAspThrLys----- 507
Db 223 GCTGCAGATTTTATGATGATGTAACCTGTGAGCTGGCTGTCTCCCCCATCCACTGCTTG 282
Qy 508 GlnAlaSerProLeuIleSerProLeuLysAsp-----Gln 520
Db 283 AACTTCATGCTGTGCTCAT-----CTCATTGAAGCCCTGAGACTAGCTTGGAGATGTTG 339
Qy 521 AlaCysProArgThrAspAspGluAspGluArgArgLysArgPheProThrAspLys 540
Db 340 GCACCTCTCTCAG-----GAAAGAGAGAGCCCTCGAGCAGAGTCCCTGGCCCAACA 390
Qy 541 AlaThrPheIleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAspLeuGlu 560
Db 391 GCT-----ACCTACATAAAGAG----- 408
Qy 561 ValIleThrSerTrpPheGlnSerThrValSerLysGluAspAlaMetProGlnAlaLeu 580

Db 409 -----TGTGTTGAGGACAGCTTGCCAGGTAATAC----- 441
Qy 581 LysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAsnPheLeuLys 600
Db 442 -----CAGGCTGCTGTAGTAAT-----GAA 462
Qy 601 GluIleGlnArgLeuAlaLeuTrpGluLysArgSerAsnAlaGlnIleArgAspTyr 620
Db 463 ACCTACCGAAGAACCTCTGGCAGCGCTAGAGAGATAAAGAGTCCCATCTACAGGTG 522
Qy 621 GlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLys----- 635
Db 523 AGCGTCTCAGATCAAGTG-----GAAGCCCAAGAGAGAGATTCGGGACCTTGAA 576
Qy 636 ---HisLeuAlaAlaHisLeuTrpLysHisSerGluAlaLeuGluAlaLeuGluAsnGly 654
Db 577 GTGTGTCTGGAAGGCCACAGGTGAAGCTCAATCTGCAGAGAGATGCTTCAGCAGAG 636
Qy 655 IleLysSerSerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnLysValCys 674
Db 637 CTGCTAAGTCCGACATCTCTGGAAGACCAAGATTGATCTG---ATGACTGAGGTCT 693
Qy 675 TyrLeuProLeuAsnThrPheLeuLysArgPro----- 685
Db 694 GAGCTGAAGCTCAAGCTGTTGTGTATGGAAGAAAGACAGAAAGAACAGAAAAACAG 753
Qy 686 -----LeuHisArgLeuMetHisTyrLysGlnValLeuGluArgLeu 699
Db 754 AGAAAGACAGAGAGATTACTGCAAGAGCTTAAAGCACTCAAAATCAAGGTGAGAGCTG 813
Qy 700 CysLysHisHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGlu 719
Db 814 GAGAACGAGCGGAACCAATAGATGAGTGAAGCTGAAGCCACCAAGCTGAGGTAGCCAG 873
Qy 720 IleThrGluMetValAla-----GlnLeuHisGlyThrMetIle 732
Db 874 CTGCAGAAACAGAGTGGCCCTTAAAGATGACAGAAATTGAGCGCTCCACAGCAGCTGCC 933
Qy 733 Lys-----MetGluAsnPheGlnLysLeuHisGluLeuLys 744
Db 934 CGGAGTGCAGCTTCCACAGCAGCATGACAGAGAGATCAAGAAATCCACCGTCTGAAA 993
Qy 745 LysAspLeuIleGlyIleAspAsnLeuValProGlyArgGlu---PheIleArgLeu 763
Db 994 -----ArgGGAATGGAACACTGCTGTTGCCATGAGATTAAGACCGTCCGATA 1044
Qy 764 GlySerLeuSerLysLeuSerGlyLysGlyLeuGlnArgMetPhePheLeuPheAsn 783
Db 1045 GAGGAGCTGACAGGCGTGTGAACAATTAAGGGTAAAG----- 1086
Qy 784 AspValLeuLeuTyrThr-----SerArgGlyLeuThrAlaSerAsnGlnPheLys 800
Db 1087 GAGATTTGATGGAGCACTCAGGCGCTTCACAAAAGAACCTCTCCATCAATGAAGATGAA 1146
Qy 801 ValHisGlyGlnLeuProLeuTyrGlyMetThrIleGlnGluSerGluAspGluTrpGly 820
Db 1147 ATTAAGGGAAGCTTCCGAAAATGGAATACCAATTAAGGCCAGAGAGAA----- 1197
Qy 821 ValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAlaSerSer 840
Db 1198 GTCCCG-----AAGCA----- 1209
Qy 841 ArgSerGluMetGluLysTrpValGluAspIleGlnMetAlaIleAspLeuAlaGluLys 860
Db 1210 -----GAGATTCACCAACA 1224
Qy 861 SerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLysSerProAsp 880
Db 1225 TGCAGCTCTCCACCCA-----GACCACTCT-----CCTTTG 1257
Qy 881 GluAlaThrAlaAlaAspGlnGluSerGluAspAspLeuSerAlaSerArgThrSerLeu 900

```

Db      1258 CCACAGAAATCACTGAGAGCAGGGCTCAGAGAAACTCTCTGAGTTCAGAAAGACTTG 1317
Qy      901 GUAArglnAlaProHisAryS-----GLVAnthrMetValHisValCysTrpHis 917
Db      1318 AGAGGTGAATCTGGGAGATATAGTGTGATGGAAACAGCTGTCTCCAGTGGAGAGCCC 1377
Qy      918 ArgAnThrSerValSerMetValAspPheSerIleAlaValGluAenGlnLeuSerGly 937
Db      1378 AAGGACAGCTCTTTC-----CTAGCGGAGCAGAAATACCCCC 1413
Qy      938 AsnLeu-LeuArgLysPheIysAsnSerAenGlyTrpGlnLysLeuTrpValValPheTh 957
Db      1414 ACATTACTGGGAGACTTT-----CAGAGGACACACCAATGAGAAAGCTGCC 1461
Qy      957 rAnPheCysLeuPhePhePheTyrlYsSerHisGlnAspAsnHisProLeuAlaSerLeuPr 977
Db      1433 -----CAGAGGACACACCAATGAGAAAGCTGCC 1461
Qy      977 OLeuLeuGlyTyrlSerLeuThrIleProSerGlnSerGlnAsnIleGln 993
Db      1462 AAATCTC-----CTCCACACTGCCTCCTCCAGGCTGACTTCCAG 1501

```

RESULT 9

```

US-10-934-272A-1
; Sequence 1, Application US/10934272A
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Sasetti, Christopher M.
; TITLE OF INVENTION: Selectin Ligands
; FILE REFERENCE: UCAL-097CON2
; CURRENT APPLICATION NUMBER: US/10/934, 272A
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 09/860,298
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 09/243,560
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: US 60/074,389
; PRIOR FILING DATE: 1998-02-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-934-272A-1

```

Alignment Scores:

```

Pred. No.:      3.69      Length:      2269
Score:          131.00     Matches:      135
Percent Similarity: 32.3%   Conservative: 68
Best Local Similarity: 21.5% Mismatches:    211
Query Match:     2.4%      Indels:      216
DB:              6        Gaps:       29

```

US-09-555-342B-2 (1-1045) x US-10-934-272A-1 (1-2269)

```

Qy      5 GlnGlnArgPro-ThrProGlySerArgLeuGlyAlaProGlnAsnSerGly----- 21
Db      647 CAGCAGACGCGACAGCCCGAGGGGCCACCAAAAGACGATGAAAGACTTCGGGACACGAGCC 706
Qy      22 -IleSerThrLeuGlnArgLysGlnLysProProProThr-----ProSerGlyLys 38
Db      707 CTCATCAGTGTGGAGGTGGAGACAGACATGGGGCCAGCTGTGCTGCTTCAGTCAC 766
Qy      38 bLeuValSerIleLysAlaGlnMetLeuAspAspThr---GlnGlnAla----- 53
Db      767 CCCAACTACAGTACTCCGGGGGACACAGACTCCACAGCCAAAGGCAAGAGCCACAGT 826
Qy      54 -----PheGlnValProGlnArgAlaProGlyLysVal 64
Db      827 GCTGCCACGTCGACGGGCTTGGGGGTGAGAGTTGAGAGCTCTTCAGGAAGACCAAGCAAGC 886
Qy      64 IleuLeuAspAlaValCysAsnHisLeuAsnLeuValGlnGlyAspTrp----- 80

```

```

Db      887 CACTGCAGAGCAGCTGT-----TTGTCTGGCCAGCAGCAGAGAGTCCC 931
Qy      81 -----PheGlnLeuGlnPheProAspHisLys 89
Db      932 GGCCTTGCTTCATTCCTCTCAAAACACAGCTCCCAAGTGGGGCCGACCAATGAGAGA 991
Qy      89 sLysIleThrValTrpLeuAspLeuLeuLysProIleValIleGlnIleArgArgProLys 109
Db      992 TCCCTTGCTCTAGAAACTCAGCTCTTCCCACTGGCCCTGGA----- 1037
Qy      109 sHisValValValPhePheValValLysPhePheProProAspHisThrGlnLeuGlnI 129
Db      1038 -----GACATGGAACAGACACTTCTCTGCTACCTTGGGACAGAGA 1078
Qy      129 uGlnLeuThrArgTyrlLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnIArgLys 149
Db      1079 AGATCTAACACAGACTCTTAGAAGGGCAGCGACTGA-----GCTCAATCCAGGAT 1132
Qy      149 u-----ThrCysAsnAspThrSerAlaAlaLeuLeuIleSerHis 162
Db      1133 ACCCTGGAGATTCTACGAGGTGATCTGCAGAGACTGGAGCAATCTGGCTGGAAAACTA 1192
Qy      162 sIleValGlnSerGlnIleGlyAspPheAspGlnAlaLeuAspArgGlnHisLeuAlaLys 182
Db      1193 CATCATCTGACATGACATGACAGAGAAACATAGACTGTGAGTGTTCGCGCAGACCGGGGG-- 1250
Qy      182 sAsnLysTyrlIleProGlnIleAspAlaLeuGlnAspLysIleVal-----G1 198
Db      1251 -----CACAGCTCTGTGGCTGGCGGTGGAGAGAGTGTCCGCCCATGGCAG 1297
Qy      198 uPheHisAsnAsn-----HisIleGly--GlnThrProAlaGlnSerAspPheGlnLe 215
Db      1298 TGGCAGACCATGGGGCCCTGGCAGACATCTCTGAGCAGAGCCAGGAGAGAGAGCAGCT 1357
Qy      215 uLeu----- 216
Db      1358 TCTCATGACACGTGTGGCGGAGGAGGGGTGTGCCCACTCAAGATGCTCTTCCATGCT 1417
Qy      217 -----GlnIleAlaArgArgLeuGlnMetTyrlGlyIleArgLysHisProAlaLysAspAr 235
Db      1418 GGGTGACATCGCAGAGGCTGGAGAGATTTGGCACTGCAGAACTATTCACCAACACGACAG 1477
Qy      235 gGlnGlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPh 255
Db      1478 CTGCCAGAGCGCGGGCCAGACGAGGTGGCGCAGGCACTACGCAAGCTC---TTCCGTGCT 1534
Qy      255 eThrLysIleAsnAlaPheAsnTrpAlaLysValArgLysLeuSerPheLysArgLysAr 275
Db      1535 GGTGTGCTATTGGGGCC----- 1550
Qy      275 gPheLeuIleLysLeuArgProAspAlaAsnSerAlaTyrlGlnAspThrLeuGlnPheLe 295
Db      1550 ----- 1550
Qy      295 uMetAlaSerArgAspPheCysLysSerPheTrpLysIleCysValGlnHisAlaPhe 315
Db      1551 -----AATGTGATGATCAT-CATCATTTGGCG 1572
Qy      315 ePhe-----ArgLeuPheGlnLysProLys-ProLysProLysProValLeuPheS 332
Db      1573 TTGGCTGTCTTCAACTGTCTGGCAGCGCGGTGGCCCAAGCTCAAGCAGCAGTGT-----T 1626
Qy      332 eArgGlySerSerPheArgPhe-----SerGlnArgThrGlnLysGlnValLeuAspTr 350
Db      1627 GGCACGGCAGAGAGCTGGCTGTGTGGAGAACGGCTGCCACGACCAACCCACAGCTGGACG 1686
Qy      350 yValLysGlnGlyGlnHisLysValGlnPheGlnArgLysHisSerLysIleHis 370
Db      1687 TGGCAGCGAC-----AGCCAGTCGAGATGAGAGAGGACCCACAGCTGAAGC 1737
Qy      370 eTrlLeaSerLeu-----AlaSerGlnProThrGlnLeuAsn 383

```

```

Db      1738 GCGGCGGGCCCTCAAGCGCCCGGAGAGCTGGGGGGCTCTATGGGGGCAAGCGGAGCC 1797
Qy      383 eRgIValleuGlInserGInserThSerLeuThrPheIylGluValaGluS 403
Db      1798 CCGAGACTCGGACTGTTTCAGAGAGACACGCACCTGTAGCGGAGCGAGCGGAGGCC 1857
Qy      403 eRProGlyGln-serCyAArgRgIylYlGluProlYsValSerIaGlyGluPro 422
Db      1858 GAGTGGCGCCGACGAGACCAAGAGAGTGA-----CCCCA-----AACGAGCGGCC 1905
Qy      423 GlySerHisProSerProIaProIaRgRgSerProIaGlyAsnLYeGlnIaAspIy 442
Db      1906 GGAAGCCGACACAGCCCGCGCTACCGCGCGCCCG-----CGCTGGCCCTCGGC 1959
Qy      443 AlaIaSerAlaProThGluGluGluValValIyAspArgThGInGInser 462
Db      1960 GCGGAGCTCTTCCCGCTT----- 1977
Qy      463 LysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHleuSerGlu 482
Db      1978 -----CCCCC----- 1983
Qy      483 LeuSerValaAsnSerGInGlyValaIaProIaAsnValThrLeuSerProAsnLeu 502
Db      1984 -----ACTTCACACGGCGGCTTCGGAACA-----ACTCCCTCAGCTC 2019
Qy      503 SerProAspThrIyGlnIaIaSerProIeuleSerProIeuleAsnAsp---GlnIa 521
Db      2020 CCGGCGCGGAGGAGCGCTCAAGCCCGCTTGCGCCGCTTCCCGCCCTGAACCCCG 2079
Qy      522 CySProArgThr-----AAspArgIuAspGluGlyArg 533
Db      2080 GCCCGCGCGGGCGGCGCGGCGCTTCTGCGCCCGGAGCTCAATTAAACCGCGGAGA 2139
Qy      534 LysArgPheProThrAspLys 540
Db      2140 CCACGCGGGCCGACGAGAAA 2160

RESULT 10
US-10-529-655-20
; Sequence 20, Application US/10529655
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research, et al.
; TITLE OR INVENTION: Human Sarcoma-Associated Antigens
; FILE REFERENCE: L0461.70156M00
; CURRENT APPLICATION NUMBER: US/10/529,655
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 2676
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-529-655-20

Alignment Scores:
Pred. No.: 5.17 Length: 2676
Score: 130.00 Matches: 190
Percent Similarity: 33.4% Conservative: 118
Best Local Similarity: 20.6% Mismatches: 374
Query Match: 2.4% Indels: 244
DB: 6 Gaps: 39

US-09-555-342B-2 (1-1045) x US-10-529-655-20 (1-2676)
Qy      137 AlaLeuGlnValIyGlnAspLeuIaGlnIyArgLeuThrCyAsnAspThrSerAla 156
Db      48 GCGCTCCAAAGTTGGAGAGAGCTG---CAGGCGCGCTGGCGCGAACCAGACCTTAAG 104
Qy      157 AlaLeuLeu-----IleSerHisIleVal 164
Db      105 AAGCATTAATAATATTGAAGAACTTCACCTGCTATTATACAGTAGACATCTTTCG 164

```

```

Qy      165 GluSerGluIleGlyAspPheAspGluIaLeuAspArgIuIleLeuAlaYsAsnLYs 184
Db      165 GAGACTGGGGTTGGGAAAACAGTAAATAGCTTG---CGAAAACAGACATGTTGAGAC 221
Qy      185 TyrIleProGlnIaAspAlaLeuGluIaAspIyIleValIgluPheHisIaAsnHisIle 204
Db      222 TTGGCAGAGGACCTAGTGGCCAGAGAAAGCTGTTCTCTGTCGAAAGAAAT----- 275
Qy      205 GlyGlnThrProIaGluIleSerAspPheGlnIleuLeuGluIleAlaArg 221
Db      276 ---GCTGAGCTGTATGAACAGACTTTGAGAAAGCAATTCGCCAAACCGCTCGGAT 332
Qy      222 -----LeuGluIle-TyrGlyIleArgLeuHisProAlaIyAspArgGly 236
Db      333 GCCCTGCAAGAGAGAGAGATGAGGGGAGACTACCAAGAACTTGAAAGCCAGGGG 392
Qy      226 uGlyThrIyHisIleAsnLeuAlaValaIaAsnThrGlyIleLeuValPheGlnIyPheTh 256
Db      393 AGCGATCTCTATAGCCCTGACACAGCAGCAAGAAACATAGG-AAACTCGGAGCTCGA 451
Qy      256 rIyIleAsnAlaPheAsnThrAlaIyValaIArgLYeLeuSerPheIyAspGlyArgPh 276
Db      452 GAGACCTCACAAGTGTCTCAGCTCATGAGAGAGATGAGAGAAAGAGGTGTAC- 509
Qy      276 eIeuIleIyAsnArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeu 296
Db      510 -----AGAAATGCACCACTTACTCTTCAGACCCCTGAGCTTGTATAGCCAGT 562
Qy      296 tAlaSerArgAspPheCyAsnLYeSerPheThrIyIleCyAsnLYeGlnHisIaIaPhePh 316
Db      563 TCAATCCCTCCATCTTGTACAGTCTCATGATGATGATGATGATGATGATGATGATGATG 622
Qy      316 eArgLeuPheGluGluPro-----LysProlYsProIyProValIlePheSerAr 333
Db      623 GAGAGAGACAGAGAGCCCATGTTTTCACACAGAAAGCT-----GGGA 667
Qy      333 gGlySerSerPheArgPheSer-----GlyArgThrGlnIyGlnValIleuAspTy 350
Db      668 AGGCCACAGCAATGCTTTCAGACAGACTGCGGCGCCAGCCAGAACAGACACCTGGGT-- 725
Qy      350 rValIyGlnIyGlnIyIyIyIyIyValaIgluPheGluIyArgIyHisIerIyIleHisIe 370
Db      726 ---GAACCCCATGGAAGGGGTTGTGATCAAAACAAAGAGACCAATATCTCCACAA 781
Qy      370 rIleArgSerIeuleIaSerGlnProThrGluIleuAsnSerIuValleuGluGlnSerG 390
Db      782 GGAACAAAGCCCGGTGATGCCAAGATGATGAGAGAGGCTCTGTGGAGACAGAGAGA 841
Qy      390 nGlnSerThrSerLeuThrPheGlyGlnIyValaGluSerProGlyGlnSerCyAs 410
Db      842 ATCACACAAAGCCCTCTTCAAAGAGAGAACCCGACCTTCAGAGGGA-CAATGCAA 900
Qy      410 g-----ArgGlyIyGluPro-- 415
Db      901 GGGAGAAACCGCCCTAGTGCGGTAAAGAAAGAGAGAGACAGAGAGGCGAGACCTGA 960
Qy      416 ---LysValSerAlaGlyIuProGlySerHisProSerPro----- 428
Db      961 AGAAGAGTGTGTTCTCTCCAGAGGCGGCTTCAGACACACCTGTAAGAAAGCCAAAGC 1020
Qy      429 ---AlaProArgArgSerProIaGlyAsnLYeGlnIaAspGlyAlaIaSerAlaPr 447
Db      1021 ACAGAGACCCAGAAAGAACCAATTTGACAAAGACCAAGCAAGGTCTGG-ACAGCTTTGAC 1079
Qy      447 oThrGluGluGluGluVal-Val 456
Db      1080 ACAGAAAGAGAGAGAGAGAGCTTGTGCCAAGTTAAAGAGAGAGGTTCTTAACACCTA 1139
Qy      456 yAspArgThrGlnGlnSer-LysProGlnProProGlnProSerThrGlySerLeuThr 475
Db      1140 AAGACTCCAGAAAGAAAGTAAATTAATTGATTAAGAACTACGCGGCTCCCTC-- 1196
Qy      476 GlySerProHisLeuSerGluIleuSerValaAsnSerGlnIyGlyValaIaProIaAsn 495

```

```

Db      1197  -----CTTAAGTTGAGAGACAGATATGAGAGATGAA-----TTCGAGCAGCCAAACC 1244
Qy      496  ValThrLeuSerProAlaMetLeuSerProAsp-----ThrValGlnAlaSer 510
Db      1245  ATGCTTTTGAATCTTACCTGATGATGACCAAGCCCGGAGAGAAAAAGAAATGTTGTG 1304
Qy      511  ProLeuLeuSerProLeuLeuAspGlnAlaCysProArgThrAspAspGluAspGlu 530
Db      1305  AAAAATTACAGCAGGACCTTGAGATTAAGACTTAAAAAATGACTCTTAAAAAGCACT 1364
Qy      531  G1yArgArgGlyValArgPheProThrAspIleAlaTyrPheIleAlaValGlyValSerThr 550
Db      1365  GGTTAAAAAC-----TTGGACTCAGATTGAGAAATTAACCAAGAGTGAACAAAAACC 1412
Qy      551  ----ThrGluArgThrTyrIleLeuValAspLeuGluValIleThrSerTyrPheGlnSerThr 569
Db      1413  AAGTCAGAGAAAGCCGGCTGAGCTGATTTA----- 1442
Qy      570  ValSerIleValAspAlaMetProGluAlaLeuValSerLeu----- 583
Db      1443  ---GCCAAGCTGAGAAAGAGTCCCTGATGTGTGGCAGTGTGGCAGACCTCCCGTTACC 1499
Qy      584  ----- 1500
Db      1500  GCGATACAGGCCAATTAACCTGCACTGCTTCCTCGAGCTGATATCTCTCTCCAGCCCA 1559
Qy      591  LeuHisIleYsrPheHisThrAsnPheLeuValGluIleGluGlnArgLeuAlaLeuTrpGlu 610
Db      1560  AAGCGAAAGAGCTTCTCTTCA-----CCCGAGAGAAAGAGAGCTGATTACTGGG 1613
Qy      611  G1yArgSerAsnAlaGlnIleArgAspTyrGlnArgIleGlyAspValMetLeuValAsn 630
Db      1614  CGCGAATGAAATTCGAAGATCCAGCTGAT----- 1643
Qy      631  IleGlnIleMetIleYsr-----HisLeuAlaAlaHisLeuTrpValHisSerGluAlaLeu 648
Db      1644  ---TCTGGTTCCAAAGTGTGCTATATCTCCCTAAATGATGACCTTGACACGACATGATC 1700
Qy      649  G1uAlaLeuGluValGlnGlyIleYsrSerSerArgArgLeuGluValAsnPheCysArgAspPhe 666
Db      1701  CGAGTACTTAAAAACAATCATGATTCATATC-----TTT 1733
Qy      669  GluLeuGlnIleValCysTyrIleYsrProLeuValHisThrPheLeuLeuArg----- 684
Db      1734  GAAGTGGAGAGAGTCCCATCTCTGTTCTTGAACCCGTTTGGAGAGGTGACACTGAT 1793
Qy      685  ProLeuHisArgLeuMetHisTyrIleValGlnValLeu-----GluArgLeuCys 700
Db      1794  CAGCTGATCGCATAGAGAAATCAATCATGATTAATTGAAGAAACAGATCAATATTGG 1853
Qy      701  LysHisHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIle 720
Db      1854  AAAGTTTAT-----TCTCACCGAGACTTTAAGAGAAAGAACCCGAGAGATGATGCTG 1907
Qy      721  ----ThrGluMetValAlaGlnLeuHisGlyThrMetIleValMetGluValAsnPheGlnLys 739
Db      1908  TGGGGAAGAGATGATCACTGCGGCTTCAGAGACCC-----CGAGAGCAGCGG 1952
Qy      740  LeuHisGluLeuValYsrAspLeuIleGlyIleAspAsnLeuValProGlyArgGlu 759
Db      1953  CTACGAGTACTTAACAAGAAAT----- 1973
Qy      760  PheIleArgLeuGlySerLeuSerIleYsrGlyLeuGlnIleArgMetPhe 779
Db      1974  ---ATTCAGTTCCGACATGCCAATAAG-----CCCAAAGGCCGACAGCAAGAAAGATGCC 2024
Qy      780  PheLeuPheAsnAspValIleLeuTyrThrSerArgGlyLeuThrAlaSerAsnGluPhe 799
Db      2025  TTTCCTC---AAGCTGTGTGGCAAGCACTCTGTACCTCCGAGAGAGGACGAAAAATTT 2081
Qy      800  LysValHisGlyGlnLeuProLeuTyrGlyMetThrIleGluGlnSerGluAspGluTrp 819

```

```

Db      2082  GGAACGGAGAGACAGCTGTCCCTGAGAAATCAAGATCAAGCAGCCCGTACCCCATG 2141
Qy      820  G1yValProHisCysLeuThrIleValArgGlyGlnArgGlnSerIleIleValAlaIleSer 839
Db      2142  GGAAGCAGCCAT-----GTTCCGCAAGT 2165
Qy      840  SerArgSerGluMetGluValYsrTrpValGluAspIleGlnMetAlaIleAspLeuAlaGlu 859
Db      2166  AGC----- 2168
Qy      860  LysSerSerSerProAlaProGluPheLeuAlaSerSerProProAsnValYsrPro 879
Db      2169  ATCAGCTTTAATCCCGCCCTGAGAGCCCGCTTATGATGCCCAAGCAGCAGATGCC 2228
Qy      880  AspGluAlaThrAlaAlaAspGlnGluSerGluAspAspLeuSerAlaSerArgThrSer 899
Db      2229  CACTGGCACCAGTGGTGCAGACACTGTTTCTTATGAT-----CCTAGAAAACCACT 2282
Qy      900  LeuGluArgGlnAlaProHisArgGlyAsnThrMetValHisValCysTrpHisArgAsn 919
Db      2283  GTGAGAAATTTGCCCAATGATGGCCAGACAAATTAA----- 2321
Qy      920  ThrSerValSerMetValAspPheSerIleAlaValGluValGlnLeuSer----- 936
Db      2322  -----GCTTTCAAGAAACAGATTCTCCGACGATAA 2351
Qy      937  -----GlyAsnLeuLeuArg-----LysPheYsrAsnSerAsnGly 948
Db      2352  ACTGAGACTTGGCTTGAAATGAAATCTGGGAGGAGAGAAATCAAGACAGTGGGGGT 2411
Qy      949  Trp 949
Db      2412  TGG 2414

```

```

RESULT 11
US-10-576-496-45
; Sequence 45, Application US/10576496
; GENERAL INFORMATION:
; APPLICANT: Teijin Pharma Limited
; APPLICANT: Komori, Toohihisa
; APPLICANT: Kanacani, Naoko
; APPLICANT: Yoshida, Carolina Andrea
; APPLICANT: Zanna, Akira
; APPLICANT: Kobayashi, Shinji
; APPLICANT: Yamana, Kei
; TITLE OR INVENTION: BONE AND/OR JOINT-DISEASE-ASSOCIATES GENES
; FILE REFERENCE: 094468
; CURRENT APPLICATION NUMBER: US/10/576,496
; CURRENT FILING DATE: 2006-04-20
; PRIOR APPLICATION NUMBER: PCT/JP2004/015879
; PRIOR FILING DATE: 2004-10-20
; PRIOR APPLICATION NUMBER: JP 2003-359172
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45
; LENGTH: 3543
; TYPE: DNA
; ORGANISM: human
; NAME/KEY: CDS
; LOCATION: (373)..(3003)
US-10-576-496-45

Alignment Scores:
Pred. No.: 11
Score: 127.00
Percent Similarity: 33.8%
Best Local Similarity: 19.4%
Query Match: 2.3%
DB: 6
Length: 3543
Matches: 138
Conservative: 102
Mismatch: 278
Indels: 193
Gaps: 31
US-09-555-342B-2 (1-1045) x US-10-576-496-45 (1-3543)

```

385 ValLeuGlnGlnSerGlnGlnSerThrSerLeuThrPheGlyGlu----- 399
18 CTGGGAGCGGGTCCAGAGAGGCTTCTTGCGCCCGAGAGAGGAGCGGGTCCCAAGAGG 77
400 -----GlyAlaGluSerProGlyGlyGlnSerCysArgArgGlyGlyGlu---ProLys 416
78 GTGGCTCTGTGGAGAGTCCCGGGGTGGGGGGCGGGCTCCGGGAAGAAGCTTCCGAG 137
417 ValSerAlaGlyGluProGlySerHisPro----- 426
138 GTCCCCCGCCCGTCACGTGGGGCGCGCCCGCCGCTGCGGTCCGCTGGTGGTTC 197
427 -----SerProAlaProArgArgSerProAlaGlyAsn 437
198 GGGGGCTTGGTCCGCGACGTGGTGGGAGCGAGCGCCGCTGGCTCTTGGGGCTCT 257
438 Lys-----GlnAlaAspGlyValAlaAspSerAlaPro-ThrGlnGlnGluGlu---GluVal 454
258 CATGTTGAAGGTGGAGAGGACACCGAGCGCGCCGACACCTGAGCCCGCCGAGAGA 317
454 ValLysAspArgThrGlnGlnSerLysProGlnProGlnProSerThrGlySerLe 474
318 GCTTCGGCGCCGTCACCGAGTAAAGAGAGAGGAGCCAGCGCAAAAGAGTCATGAG 377
474 uThrGlySerProHisLeuSerGlnLeuSerValAsnSerGlnGlyValAlaPro-- 493
378 TTCTGATGTCTAGTACGCGCTGGAGAGCTGCCCTGAGCAAAATGACGGAGATCATTGACAG 437
494 -----AlaAsnValThrLeuSerProAsnLeuSerProAspThr 506
438 CACTAAACAGGTGACGATCTTATGATGATGATCTTGATGAGCTGAGCTGGCTCCCGCGC 497
506 rLysGlnAlaSerProLeu---LysSerProLeuLeuAsnArgGlnAlaCysProArgTh 525
498 CTCTACATGAACCCCTTCCCGGTCTCCATCTCATTCAGAGACTTGAGGCTGGCTTGG-- 555
525 rAspAspGlnuArgGlnuArgArgLysArgPheProThrAspLysAlaTyrrPheAla 545
556 -----GAGATGCTGAGCTTCTCAGAGAGAGAGAGAGCCCTCTGAG 596
545 AluGlnAlaSerThrThrGlnuArgThrTyrrLeuLysAspLeuGluValIleThrSerTr 565
597 C---CAGATCCCTGGCCCAACAGAGCTGCTACATAAGAA-----TG 635
565 pPheGlnSerThrValSerLysGlnuAspAlaMetProGlnuAlaLeuLysSerLeuIlePh 585
636 GTTGAAGAGAGCTTGTCCAGAGTAAACAC----- 666
585 eProAsnPheGlnuProLeuHisLysPheHisThrAsnPheLeuLysGlnuIleGlnuAsn 605
667 -----CAGAGTGTCTGTAGTAAAT-----GAAACCTTACAGAGAG 701
605 GLeuAlaLeuTrpGlnuArgLysSerAsnAlaGlnIleArgAspTyrrGlnuArgIleGlyAs 625
702 CTTGACAGCTGAAGGAGGATTAAGAGAGTCCCTCATATTGCAAGTGAATGCTGCACAGA 761
625 pValMetLeuLysAsnIleGlnuLysMetLys-----HisLeuAla 639
762 CCAAGTA-----GAAAGCCAGGAGGAGAAAGATTGAGAGCTGGAAGTGTCTGGAAGG 815
639 HisLeuLeuTrpLysHisSerGlnuAlaLeuGlnuAlaLeuGlnuAsnGlyIleLysSerSerAs 659
816 ACACCAAGGTGAATCTCAATGTCTGTGAAGAGAGTCTTCAACAGAGAGCTGTAAGCCGAC 875
659 gArgLeuGlnuAsnPheCysArgAspPheGlnuLeuGlnuLysValCysTyrrLeuProLeuAs 679
876 ATCTCTTGAAGACCCAGAGAGCTGATCTG---ATGACTGAAGTGTGAGCTGAAGCTGA 932
679 nThrPheLeuLeuArgPro----- 685
933 GCTGGTTGGCATGAG 992

686 ----LeuHisArgLeuMetHisTyrrLysGlnuValLeuGlnuArgLeuCysLysHisLeuPr 704
993 GTTACTGCAAGAGCTCAGAGACCTTCAAAATCAAAAGTGAAGAGTTGGAAATGAAGAGAA 1052
704 oProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGlnuIleThrGlnuMetVal 724
1053 TCAGTATGAATGAAGAGTAAAGGCCATGAAGCTGAATGCCCGCAGCTGCAAAACAGAGT 1112
724 Ala-----GlnLeuHisGlyThrMetLeu----- 733
1113 GGGCTGAAAGATGACAGAAATTAAGCTGTGACAGCCAGCTCTCCGAGACAGAGCTCT 1172
734 -----MetGluAsnPheGlnuLysLeuHisGlnuLeuLysAspLeuI 748
1173 CCACAGTAGAGTACACACAGAGAGAGACCAAGAAATTCACAGCTGGA-----AT 1223
748 eGlyTLeAspAsnLeuValValProGlyArgGlnuPheIleArgLeuGlySerLeuSerLys 768
1224 GGGAGTGAACCTTGTGCTTGCCATGAAGATGAAGACCGT-----CGGATGAAGGA 1277
768 sLeuSerGlyLysGlyLeuGlnuArgMetPhePheLeuPheAsnAspValLeuTy 788
1278 GCTTACGGGGCTGTTAAACAGTACCGGAG-----GTAAAGAGATGTGATGCT 1328
788 rThr-----SerArgGlyLeuThrAlaSerAsnGlnPheLysValHisGlyGlnLe 805
1329 CACTCAAGGGCTTCCGAGAGAACTCTCTCATCATATGAAGAAGAACCGAGAGAGCTTT 1388
805 uProLeuTyrrGlyMetThrIleGlnuLysSerGluAsp-----GluTrpGlyValPr 822
1389 CAGAGAGTGAACGCTCAAAATTAAGACCTGAGAGAAATTATTAAACAAGAGATCTTCC 1448
822 HisCys-----LeuThrLeuArg 828
1449 AAGATGTAGCTCTCTCAAGTGGGGCCACTTCATTCGCACAGAAATCACTGGAAGCCAG 1508
828 gGlyGlnuArgGlnSerIleIleValAlaAlaSerSerArgSerGlnuMetGluTyrrPva 848
1509 GGGTCAGAAAAGCTCTTGTGATGTTGAAGACTTGAAGAGTGAATGTGTGAGTAAAGT 1568
848 IGlusprIleGlnuMetAlaIleAspLeuAlaGluLysSerSerSerProAlaProGluPh 868
1569 TATGATGAGAAACAGAGCCCTTCCGGGTGTGAACCCAGAGAGAGCCCT-----TT 1619
868 eLeu----- 869
1620 CTGGCGAGAGCAAAATATCCCACTTAACTGGAGAGCTTCAAGAGCCAGCCCAATGG 1679
870 ----AlaSerSerProProAspAsnLysSerProAspGlnuAlaThrAlaAla----- 885
1680 AGAGGCTGCCAAATCTCTCCACCATCTGCGACGCTGACCCAGGGAGACAGCTGCT 1739
886 -----AspGlnuLys 895
1740 GAGGCTGAGAGACACAGAAAGTGGCTGGAGACAGACTGTGTGCAATGACCTTCAATC 1799
895 aSerArgThrSerLeuGlnuArg-----GlnAlaPro-----HisArgI 908
1800 CACATCACTGGGCACTGAATCAGGCTCTCATCTCTCTGACACCAAGTGTAAACGGAA 1859
908 yAsnThrMetValHisValCysTrpHisArg----- 918
1860 TCCCAAGGCACTTAAGAGTTCTGGGAGAAAATCCGAAGAACTCAGTACAGAAATTTCTA 1919
919 -AsnThrSerValSerMetValAspPheSerIleAlaValGlnuGlnuLeuSerGlyAs 938
1920 CACTGACACGCTGGAGATGGAGAGTTTCAAGAGAGTGGCTCCGGGCAACCGAGAGGCC 1979
938 nLeuLeuArgLysPheLysAsnSerAsnGly 948
1980 AAGACTCTTAGAGACAGAGAGCTCAAGGGA 2010

RESULT 12

```
US-60-796-903-4
; Sequence 4, Application US/60796903
; GENERAL INFORMATION:
; APPLICANT: Johnston, Patrick
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Colorectal Cancer Microarrays and Methods of Use
; FILE REFERENCE: 55815-300 (127365)
; CURRENT APPLICATION NUMBER: US/60796,903
; CURRENT FILING DATE: 2006-05-02
; NUMBER OF SEQ ID NOS: 1071
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 6373
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-60-796-903-4

Alignment Scores:
Pred. No.:      23.8      Length:      6373
Score:          126.50    Matches:      144
Percent Similarity: 37.2%  Conservative: 104
Best Local Similarity: 21.6% Mismatches:    243
Query Match:     2.3%     Indels:       177
DB:              8        Gaps:         32

US-09-555-342B-2 (1-1045) x US-60-796-903-4 (1-6373)
QY      343  ThrGlnlysgInvalleu---AspTyrVallysgGluGly----- 355
DB      897  ACACCTAAGCAGATCTCTCAATTAACCTCAAGAAAGAACACTGATATGACGATGCT 956
QY      356  ---HisIlylyValGln-----PheGluArgIlyHisIserIlyIleHiserIleArg 372
DB      957  ACTAAAGAAAGATCAACTTGCTCTTACG-----AAATCCAGACCTTAAAG 1004
QY      373  SerLeuAla-----SerGlnProThrGluLeuAsnSerGluVal 385
DB      1005  TCTCGACACAGCTGGAGCGCCCAAGAAATACCAAGCTTAAATCCACATCTGAAGTC 1064
QY      386  LeuGlnIlySerGlnIlySerThrSerLeuThrPheGluGly-----Ala 401
DB      1065  AAAATCTATTGTAACAGAAAGAAAGCAAGTTGACATAGA-AGTGGCTGAATTCAGAGACA 1123
QY      402  GluSerProGlyGlyGlnSerCys-----ArgArgGlyLyseGluProLyVal 417
DB      1124  GCTTCAACTGGAAGTCAAGATCAACAGACATTTAAAGAGAGAGAGAGAGAGAG--- 1180
QY      418  SerAlaGlyGlnProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsn 437
DB      1181  -----AGCAAACTTGAAGAGCTCCG----- 1201
QY      438  LyseGlnAlaAspGlyAlaAlaSerAlaProThrGlu-----GluGlnGluGluVal 454
DB      1202  -----AAGCCAAACACACGAAAGGTGAGAGAACTCCAC 1237
QY      454  IValIyAspArgThrGlnIlySerIlyProGlnProGlnProSerThrGlySerIle 474
DB      1238  ATTGCACACACGACTGGAAGAAAGTGAAGGGAGCT-CCGAAAGAACTGGAGAGACTAT 1296
QY      474  uThrGlySerProHisIleuSerGluLeuSerValAsnSerGlnIlyGlyValAlaProAl 494
DB      1297  TCCAGGTGAAGATGGAACGGAGACAGCATCAGACTGAGATCAGGAGATCTCCAGAGCCAGC 1356
QY      494  a-AsnVal-----ThrIleuSerProAsnLeuSerProArgThrIlyseGlnAlaSerP 511
DB      1357  TCTCAAGAAATGCACAGATGAACCTGACAGTGCAGAAACCATTCGAGAGACGAGAGAGAG 1416
QY      511  rIleuIlySerProLeu-----LeuAsnAspGlnAlaCysProArgT 525
DB      1417  CTCCTGATTGAGAGACTCTTACAGGCAAAACAGATCTTCAAGATCTGCTGATTCGCAAG 1476
QY      525  hAspAspGluAspGluGlyArg--ArgIlyArgPheProThrAspIlyValArgPheI 544
```

```
DB      1477  AGAGCAAGAAAGACCTTGTGAAGAAAGGAGCGTAACTCACCGCCCTGAAGGAGGCC 1536
QY      544  leAlaIyGlnValIserThrThrGluArgThrIyr----- 555
DB      1537  TGAAGAAAGAGTTTCCAGCATGATGAGAGATGGAACAAGCTGAAGAGCAATATGATG 1596
QY      556  -----LeuIyAspLeuGluValIleT 563
DB      1597  CTGAGTTGACAGCCCTTGAAGAGAGTGTGAAGAAAGCAACCAAGAAATGTGAGGCTTGG 1656
QY      563  hSerTrpPheGlnIserThrValSerIyGluAsp--AlaMetProGluAlaLeuIyys 582
DB      1657  CGAGC-----AGAGACACACTTCAGAGCAAGACCCAGCGGAGCTGAATGCGCGCTGA 1710
QY      582  eIleuIlePheProAsnPheGlnProIleuHisIlyysPheHisThrAsnPheIyAspGlu 602
DB      1711  AGCTTCTGCAGAGAGAGATGAGAGCTGCAG-----GGAAGAAAGCAAGAGC 1758
QY      602  leGlnIlyArgLeuAlaLeuTrpGluIyArgSerAlaAlaGlnIleArgAspTyrGlna 622
DB      1759  TGAAGCGGAGAGATTGCT-----CAGCTTCAAGGACAGATGAGAGACTGAA- 1805
QY      622  rGlyIleGlyAsp-----ValMetLeuIyAsnIleGlnGly----- 633
DB      1806  -----GGCGATGAAGCCCAAGCGCAAGAGAAACGCTGAAGAAATGACAGAGAAATACGAC 1860
QY      634  --MetIyHisIleAlaAlaIleLeuTrpIyHisIserGlu-----AlaLeuG 649
DB      1861  AGTTAGAGAGAGCCCTTGTGACCGCCAGAAAGAAAGAAAGAAAGAAAGAAAGCTGTGTACGCCAGAA 1920
QY      649  IuAlaLeuGluAsnGlyIleIySerSer----- 658
DB      1921  GGGCCCTGGAAGATTAAGTCAAGAGCTGTCAGAGGAATCTGAGTCACTATCCAGAGAGC 1980
QY      659  --ArgArgLeuGluAsnPheCysArgAspPheGluGlnIlyValCysTrpIleuProL 678
DB      1981  AGAAGCAGTGTGTGAGMACTCAAGAGAGAGAGAGAGAGAGAGAG----- 2024
QY      678  euAsnThrPheLeuLeuAspArgProLeuHisArgLeu-----M 690
DB      2025  -----GACACGCTAAGAGAGTTGAAGAAACAGATGAGAGATGACGCGT 2067
QY      690  eHisTrpIyGlnIlyValLeuGluArgIyCysIyHisIserProSerHisIleAspR 710
DB      2068  GGACCTCGGGCAAAACCATTGAGAAATGCAAG-----GAGATGSCACACA 2115
QY      710  heArgAspCysArgAlaAlaLeuAlaGluIleThrGluMetValAlaGlnIleuHisIe 730
DB      2116  TT-----GTTGAGGCGCTCCCGTACCTCAACCTGTGAGCTCCAGAACCAAGCTG-----G 2163
QY      730  hMetIleIyMetGluAsnPheGlnIlyLeuIyAspGluIyValIyAspLeuIleGlyI 750
DB      2164  ATGAGTTTAAGAGAAAGAAACCGCAGGAGCTCCGAGAAATGCAAGACAGCTTG----- 2216
QY      750  leAspAsnLeuValIalProGlyArgIyGluPheIleArgLeuGlySerLeuSerIyAsp 770
DB      2217  -----AAGGAGAAACGCTGAGAGGCAAGAAAGTCCGACTGA 2253
QY      770  eArgIlyGlyLeuGlnIlyArgMetPhePheLeuPheAsnAspValIleuLeuTrpHis 790
DB      2254  CAGCGATGAAGAAATGCAGAGATGAGATGCGTGTGAGAGAAAGTTACGGAGACTTACAGA 2313
QY      790  eArgGlyLeuThrAlaSerAsnGlnIyPheIyValIyGlnIlyLeuProLeuTrpIyGlyM 810
DB      2314  GAGCT-----CAGGATGAAGCACTCACAAAGAGCAGCTTGTGAGAC 2355
QY      810  eThrIleGlnIySerGluAspGluTrpGlyValProHisCysIyLeuThrIleuArgIyG 830
DB      2356  AGACGCTGAAGACCTGAGATATGAGCTGGAACCAAGACT-----CAGCTCAAGATG 2409
QY      830  IuArgIlySerIleIleValAlaAlaIserSerArgSerGluMetGluIyTrpValGlu 850
```



```
Db 2410 ACCGACGAGCTGCTC-----AGCAGATGAGGACAAGCTGCTC 2451
Qy 850 epIleGImeTAlaIleApleuAlaGluYseSerSerProAlaProGluPheLeuA 870
Db 2452 AACTGGAGATGAACTGGAGAAGAGAGAAACACTCAGATTTCCTGCTGAGAGAA 2511
Qy 870 lAserSerProApeAntySerProAepGluAlaThraAlaAlaepGluInuSerg 890
Db 2512 GTAGAGACGAGGAACACATGAGCAGCTGAGAGATGACTTCTAGAGAGAGCTGCGA 2571
Qy 890 lueApeApeLeuSerAlaSerAgtThrSerLeuGluAgtGlnAlaProHisArgGlyAnt 910
Db 2572 GACAAGACTTGAAGTGGACACAAGATTTCTCTGAGAGGACAGAACAGACTTAAGAGACC 2631
Qy 910 hmeTvalHleval 914
Db 2632 GGATTATCCACCTG 2645

RESULT 13
US-11-366-965-1/c
Sequence 1, Application US/11366965
GENERAL INFORMATION:
APPLICANT: Griffiths, Remy
APPLICANT: Holseth, Susan K.
APPLICANT: Zagursky, Robert John
APPLICANT: Mercalf, Benjamin J.
APPLICANT: Peek, Joel A.
APPLICANT: Sankaran, Banumathi
APPLICANT: Fletcher, Leah Diane
TITLE OF INVENTION: CHLAMYDIA TRACHOMATIS POLYNUCLEOTIDES AND VECTORS, RECOMBINANT HC
FILE REFERENCE: GEN-T109X
CURRENT APPLICATION NUMBER: US/11/366,965
CURRENT FILING DATE: 2006-03-02
PRIOR APPLICATION NUMBER: US/09/201,228
PRIOR FILING DATE: 1998-11-30
PRIOR APPLICATION NUMBER: US 60/107,077
PRIOR FILING DATE: 1998-11-04
PRIOR APPLICATION NUMBER: FR 97-16034
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: FR 97-15041
PRIOR FILING DATE: 1997-11-28
NUMBER OF SEQ ID NOS: 5982
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1038608
TYPE: DNA
ORGANISM: Chlamydia trachomatis
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(1038608)
OTHER INFORMATION: n = A,T,C or G
US-11-366-965-1

Alignment Scores:
Pred. No. 1 11e+04 Length: 1038608
Score: 126.00 Matches: 157
Percent Similarity: 35.1% Conservative: 108
Best Local Similarity: 20.8% Mismatches: 269
Query Match: 2.3% Indels: 221
DB: 7 Gaps: 38

US-09-555-342b-2 (1-1045) x US-11-366-965-1 (1-1038608)
Qy 324 ProlyaPro-----LysProvalLeuPheSerArgGly 334
Db 752301 CCGAGCCGGGTTTCTGTGATGAAACAGTATCACTGCCCTGTTTGTAAAGATGGG 752242
Qy 335 SerSerPhe-----ArgPheSerGlyArgThrGlnLeuSer 346
Db 752241 GTTTCCTTTGTAATGGAAGCTGCCAAGCTCTGTTTTTAAAGGGAAGAGTAGAAGCA 752182
Qy 347 -----ValLeuAapTyValIlyeGluGlyHisIlys----- 357
```

```
Db 752181 GATATAGAGTATATGCTTCACAGACAGAGCTGTATATAGGAGAGCTTCAGCGCAAGGC 752122
Qy 358 -----LysValGlnPheGlu-----ArgIlyHisSerLysIleHisSerIleArgSer 373
Db 752121 TCTTCAGCGTTCCTCCGAGAGTCTCTAAAGAGAGAGAAATCCAAAGTCTCTCC 752062
Qy 374 LeuAlaSerGlnProThrGlnLeuAmsSerGluValLeuGluGlnSerGlnGlnSerThr 393
Db 752061 GAGACAAAGTTTCTGAGAGAGTATCTAGTTTCAGCAATAGCGAAAGAAAGAGTTAGCG 752002
Qy 394 SerLeuThrPheGlyGluGlyAlaGluSerProGlyGlyGlnSerCysArgArgGlyLys 413
Db 752001 GTGTTTTTTTAGCTTCCTGTTGAGAGAGAGCTCGA-----ACACCAAAAGAACTCTCT 751948
Qy 414 GluProLysValSerAlaGlyGluProGlySerHisPProSer----- 427
Db 751947 GAGCCAAAGTCTCTCCAAAGAA-----GACACAGACTCTCTGTTCAGAGAAAAAAG 751891
Qy 428 -----ProAlaProArgArgSerProAlaGlyAmsIlyGlnAlaAepGlyAlaAlaSer 445
Db 751890 GATCTGAGCTTCCTTGGCAAGTCCAGACAACTTAACAACTATCTCCATCCAGCGAGT 751831
Qy 446 Ala---ProThrGluGluGluGlu---GluValIallyAAsPArgThrGlnGlnSerLys 463
Db 751830 GGTGAACCAACCAATCTCAAAACGCGAGTATGAGAGAAACAGAACATCCGCCGATCA 751771
Qy 464 ProGlnProProGlnProSerThrGlySerLeuThrGly-----SerProHisLeuSer 481
Db 751770 AATCAGCGCCACAGCTTCTTCGCTTCCAAATCGGGTTCAAAGTCCGAAAAATCG 751711
Qy 482 GluLeuSerValAmsSerGlnGlyGlyValAlaProAlaAmsValThrLeuSer----- 499
Db 751710 GAGCAACAACCTTCAAAAGCCTCCCAATCCCGGAAACTCCAGAGCCGTCAGAGAA 751651
Qy 500 ProAmsLeuSerProAepThrLysGlnAlaSerProLeuIleSerProLeuLeuAmsP 519
Db 751650 CTTAATAGCGGTACGAGAAAGAAATCGCGCATCTCCAAATGAGAAAGCTCCGTAACAGA 751591
Qy 520 GlnAlaCysProArgThrAspAspGluAsp--GluGlyArgArgLysArgPheProThr 538
Db 751590 GAAGCAGACTCAGGACGACGAGGAAGAAAGAGAGGCT----- 751552
Qy 539 AspLysAlaTyrrPheIleAlaLysGluValSerThrThrGluArgThrTyrrLeuLysAsp 558
Db 751551 -----GAAGAGATACTGCTGAAGCGCAGCAAAATGAAG 751516
Qy 559 LeuGluValIleThrSerThrPheGlnSerThrValSerLysGluAAsPAlaMetProGlu 578
Db 751515 CTAAAGCAGAGGCTTCT-----CAAGAAAGAGAAAGAAAGAGAAATGAANNAG 751462
Qy 579 AlaLeuLysSerLeuIlePheProAmsPheGluProLeuHisIlyPheHisThrAms-- 597
Db 751461 GTTCTGCTCCCTT-----AATGTCAGATCTTTCCGTTTGATCAAGAAATC 751411
Qy 598 PheLeuLysGluIleGluGln----- 604
Db 751410 TTCCTGCTGATAGAAAGATCTTGACAGAAACAGTGGCGTTGATTGACGCAACA 751351
Qy 605 -----ArgLeuAlaLeuTrp 609
Db 751350 TCACGATTTTGTGAAGTTCTTGCTGTCGGAATATCGGTGCTGAATTCATTTGGAT 751291
Qy 610 GluGlyArgSerAmsAlaGlnIleArgAspTyrrGlnArgIleGlyAAsPValMetLeuLys 629
Db 751290 AGTGGAAACCTATATCTATGTAAGATGATCCGAG--GTTGAGACATGCTTTAAGT 751234
Qy 630 AmsIle-----GlnGlyMetLysHisLeuAlaHisIleLeuTrpLysHisSer 645
Db 751233 GATATAGATATTTGCGCGCAACATCGCAAGATCATTTAGCGCAAT-----GATTAATCA 751180
Qy 646 GluAlaLeuGluAlaLeu-----GluAmsGly--IleLysSerSerArgArgLeuGlu 662
```


Db 4150 CTTTCCCGAGACCTGAGTGAATGCCCTTTTGCAGATCTCAGTAAATCCCTTACC 4209
Qy 469 ProserThrglySerLeuThrglySerProHileuSerGlu-----LeuSerValaen 486
Db 4210 CCAAGCCTCGACGACATGACACTTTCTCCAGACCTTGGTGAACAGATCTTTCCCAAC 4269
Qy 487 ---SerGlnGlyValaIlePro-----AlaenValaIleuSerProAsnLeuSer 503
Db 4270 TTTGGTCAGATGATCCCTTTCCCGAGACCTCAGCCAGGTGATCTTCTCCAGACATCAGT 4329
Qy 504 -----ProAspThrglySerGlnaIleSerPro----- 511
Db 4330 GACACACCTTCTCCCGGATCTCAGCAATATCAGTCTCCGACACTTGCAGATCAGATA 4389
Qy 512 -----LeuIle----- 513
Db 4390 TTTCACCTTCTGAACTTATGACATGATGCTTCTTCAAGATTATGATGATCTTTTCT 4449
Qy 514 -----SerProLeuLeuAsnAsp----- 519
Db 4450 TATCCAGACCTTGGTCAGATGCCATCTCTTCATCTCTCACTCTCAATGATATCTTTCTA 4509
Qy 519 ----- 519
Db 4510 TCAAGGAATTATTCACCTGTTATAGTGGCCCTCAGTAAAGATGTACAGATTACATT 4569
Qy 520 GlnaIleCyProArg-----ThraAspArgIleuArgGlyValaArgArg 535
Db 4570 GAGATCATCTCCAAAGAGAGAGTCCAGAGTCCAGATGAGTACTGATGATGATGATAT 4629
Qy 536 PheProThraArgIleValaIleThraIleValaIleValaIleSerThrglyValaThrgly 555
Db 4630 GTGCCCTATGATGACCCCTAC-----AAACCTGATGTTAGGCAAAAC 4671
Qy 556 Leu-----LysAspLeuGluValaIleThraSerThrglyPhe----- 566
Db 4672 ATCAACCTCCAGAGATCTGACAAACATTCAGATGCTGCTCCGACAGACAAATGGA 4731
Qy 566 ----- 566
Db 4732 AACAGAAATTATTCATTCCTGCTGAGAAATATCTGGAGTTATTCAGAAATTGTA 4791
Qy 567 -----GlnSerThraIleSerLeuGluAspAlaMetProGlu-----AlaLeuIleSer 582
Db 4792 CAAGGGAACAGATGATGAAGACTGTGATGATATTCAGAAAGATCCACATATATGAAGA 4851
Qy 583 LeuIlePheProAsnPheGluProIleuThraIleValaIleValaIleValaIle 600
Db 4852 GTAGTTTTTGAAGTAC-----CTCCAGACACCTTTTACCAAACTGAT 4896
Qy 601 -----GluIleGluGlnaIleValaIleuTrpGluIleValaIleSerAsnAlaGlnIle 617
Db 4897 CTTGAGAGGGAGTGTGAAGACATCTCGAATTTCTGCTCCTATATCAGAGCTGAAGTG 4956
Qy 618 ArgAspThrglyGlnaIleGluValaMetLeuValaMetIleGlnIleMetIle 636
Db 4957 GATGAT-----GTATCCAAAGTTCGTTTAAATTTTGAAGATCCAGACCCGATTTCT 5007
Qy 637 LeuAlaIleValaIleuTrpValaIleSerGluValaIleuGluValaIleValaIleVala 656
Db 5008 CTACATGCCCATGAGCTTCTTATGAAGAAATCATCATCAGAGAGAAAGATATATGAAGATGAC 5067
Qy 657 SerSerArgArgGluGluValaMetPheCyArgArgPheGluGluValaIleValaIleVala 676
Db 5068 TCTCTCT-----GAATGCTTTAAGAGATATGCTGTTCCAG----- 5103
Qy 677 ProLeuValaIleThraPheLeuValaIleGluValaMetIleValaIleValaIleVala 696
Db 5104 -----CCAAATGCACTTATCTTACCTTACCTGATGAGCCACT 5139
Qy 697 GluArgLeuCyValaIleValaIleProSerThraIleValaIleValaIleValaIleVala 715
Db 5140 GAGCGATCAGGGCCAGAAAGTCTGGGCT-----GCTGTGCGGGCTTGG 5184

Qy 716 AlaLeuAlaGluIleThrglyMetValaIleGlnIleuHileGlyThraMetIleValaMetGlu 735
Db 5185 GCCCTACTACGAGCTGTGAACCCAGAAAAGATATTCACAGGCTGATAGTCCCTC 5244
Qy 736 AsnPheGlnValaIleGluIleValaIleValaIleValaIleValaIleValaIleVala 755
Db 5245 CTATCTGCCAAAAGAAATCTACATAGAAC-----AGCAACATGCCCTGTG 5292
Qy 756 ProGlyArgGluPheIleValaIleValaIleValaIleValaIleValaIleVala 775
Db 5293 GACATGAGAGAAATTTGCTTACTA----- 5316
Qy 776 GlnArgMetPhePheLeuPheAsnAspValaIleuIleValaIleValaIleVala 788
Db 5317 -----TTTATGACCTTGTATGAAGAAAGAGAGCTATGAAAGAAAGTCCGA 5367
Qy 789 ThraSerArgGlyLeuThraIleSerAsnGlnPheValaIleGlyIleuProLeuIleVala 808
Db 5368 AGTCTTGGAGACTCAGATCTCTCAGAAATGAAATCCCATGAGTTTACAGCCATTAT 5427
Qy 809 GlyMetThra-----IleGluIleSerGluAspGluIleTrpGlyValaProHile 823
Db 5428 GCGATGATCTACAGCTTGCCTGCGCTGAATAATGTATAGCAAGAGTGGTGAGTTACAC 5487
Qy 824 CysLeuThraIleuArgGlyGlnaIleArgGlnSerIleIleValaIleAlaIleSerSerArgSerGlu 843
Db 5488 CTGCTGAACATGAGCGCTCCCAAGACATTCAGCTGTTCTTACAGGCCAACCTTG 5547
Qy 844 MetGluIleTrpValaIleuAspIleGlnMetAlaIleAspLeuAlaGluIleValaIleSerSer 863
Db 5548 CTGAAAATGCAATTAACAGACACAGTGAAGGCTGCGCCCTTGCCTGTTCAATT 5607
Qy 864 ProAlaProGluPheLeuAlaIleSerSerPro 873
Db 5608 AAAACTCTGAAATGAAGGATCAAAACT 5637

RESULT 15
PCT-US06-14800-70
; Sequence 70, Application PC/TUS0614800
; GENERAL INFORMATION:
; APPLICANT: Children's Hospital Medical Center
; APPLICANT: Wong, Hector R.
; APPLICANT: Aronow, Bruce J.
; TITLE OF INVENTION: METALLOTHIONEIN AS AN EARLY BIOMARKER
; TITLE OF INVENTION: FOR DEATH SECONDARY TO SEPTIC SHOCK AND AS A NOVEL
; FILE REFERENCE: CHM027.001VPC
; CURRENT APPLICATION NUMBER: PCT/US06/14800
; PRIOR FILING DATE: 2006-04-19
; PRIOR APPLICATION NUMBER: US 60/673,656
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 6914
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US06-14800-70

Alignment Scores:
Pred. No.: 32.4 Length: 6914
Score: 125.00 Matches: 162
Percent Similarity: 32.2% Conservative: 80
Best Local Similarity: 21.5% Mismatches: 279
Query Match: 2.3% Indels: 232
DB: 1 Gaps: 35

US-09-555-342B-2 (1-1045) x PCT-US06-14800-70 (1-6914)
Qy 285 AsnSerAlaTrpGlnaIleThraIleuGluPhe-----LeuMetAlaIleSerArgAsp 300
|||||
|||

Db 3690 AACTCAGCAGCAAAACCTCTCTCCAGACCTCAGCAGCAAGCAGCTCTCTCCAGACCTCA 3749
 Qy 301 PheCysAluSerPheThrIleCysVal-GluHisIleAlaPhePheArgLeuPheG1 320
 Db 3750 TTCAGAAACCTTTCCCGAGCCTCGGTGAGATGCCAT-----TTCCTC 3794
 Qy 320 uGUProLybPro-----LysProLybProVal---LeuPheSerArgG1 334
 Db 3795 CAGACCTCAGCCATCAAAACCTTTCTCCAGACCTCAGCCATCAAAACCTTTCTTAAGACC 3854
 Qy 334 ySerSerPheArgPheSerG1ArgThrGlnIleGlnValLeuAspGlyValIlyGlnG1 354
 Db 3855 TCAGCCAGCAAAACCTCTCTCCAGAACTCACTCA-----GACAAACCTTTCTCCAG 3905
 Qy 354 yGlyHisIleIleValGlnPheG1IleArgIleSerIleIleHisSerIleArgSerIle 374
 Db 3906 CCGCTCGGTGAGATGCCCTTTCTCCAGACCTCAGCCATCAAGCAATTTCTTAAG----- 3960
 Qy 374 uAlaSerGlnPro-ThrGlnIleuAsnSerGlnValLeu-----GluG 388
 Db 3961 ----CTTCAGCCAGCAAAACCTCTCCAGAACTCAGCCATGAGCTCTCTCCAGAAC 4016
 Qy 388 IAsnSerGlnIleSerThrSerIleuThrPheG1IleGlnIleValIleGlnSerProG1IleGlnS 408
 Db 4017 TCAGTCAAGCAAAACCTTTCCCGAGCCTCGGTGAGATGCCATTTCTCCAGACCT-CAGC 4075
 Qy 408 eCybArg-----ArgGlyIlybGlnProLybValSerAlaGlyIleGlnProGlySerH 425
 Db 4076 CATCAAAACCTTTCTTAAGCTTTCAGCCAGCAAAACCTCTCTCCAGAACTCAGTCAAAACA 4135
 Qy 425 IAsProSerProAlaProAlaArg---SerProAlaGlyAsnIlyGlnIleAlaAspGlyAla 444
 Db 4136 AACCTTTCCCGAGCCTCGGTGAGATGCCCTTTCTCCAGACCCAGCCATCAAAACCTT 4195
 Qy 444 IAsSerAlaProThrGlnIleGlnIleValIleValIyAspArgThrGlnIleSerIlyb 464
 Db 4196 TCTCTAAGCTTCAGCCAGCAAAACCT-----CTCTCAAGAACTCAGTCAAGCAAAAC 4246
 Qy 464 roGlnPro-----Pro-Gln 468
 Db 4247 CTTTCCCAAGACCTCAGTGAATGCCCTTTTCCAGATCTCAGTCAAAATTTCCCTTACC 4306
 Qy 469 ProSerThrIleuThrGlnIleSerProHisIleuSerGln-----LeuSerValAsn 486
 Db 4307 CCAAGCCTCAGCAGATGAGACTTCTCCAGACCTGCTGAGACAGATCTTTCCCAAAAC 4366
 Qy 487 ---SerGlnIleGlyValAlaPro-----AlaAsnValThrIleuSerProAsnLeuSer 503
 Db 4367 TTGGTCAAGATGTCCTTTCCCGAGCCTCAGCAGGTGAGCTCTCTCCAGACATCAGT 4426
 Qy 504 -----ProSerThrIleGlnIleAspPro----- 511
 Db 4427 GACACCAACCTTCTCCCGAGTCTCAGCCAGATACACTCTCTCCAGACCTTGATCAATA 4486
 Qy 512 -----LeuIle----- 513
 Db 4487 TTTCACCTTCTGAATCTAGTCACTCATGCTCTTCAAGAAATTAAATGAGTCTTTCT 4546
 Qy 514 -----SerProLeuLeuAsnPro----- 519
 Db 4547 TATTCAGACCTTGGTCAAGATGCCATCTCTTCATCTCTCAATGATATCTTTCTA 4606
 Qy 519 ----- 519
 Db 4607 TCAAAGAAATTAAATTCACATGCTTATAGTGGCTCAGTAAAGTGATACAGATTACATT 4666
 Qy 520 GlnAlaCysProArg-----ThrAspArgIleuAspGlnIleArgArgIlybArg 535
 Db 4667 GAGATCATTCCTCAAGGAAGAGGTCCAGAGCAGTGAAGTCAATGCTGAATGATTGATTAT 4726
 Qy 536 PheProThrAspIlybAlaIlyrPheIleAlaIlybGlnValSerThrThrGlnIlybThrTyr 555
 Db 4727 GTGCCCTATGATGACCCCTTAC-----AAACTGATGTTAAGCAAAAC 4768

Qy 556 Leu-----LysAspLeuGlnValIleThrSerThrPhe----- 566
 Db 4769 ATCACTCTCCAGAGATCTCTGACAAATTCGACAGTGAATCCTCCGACAGCAAAATGGA 4828
 Qy 566 ----- 566
 Db 4829 AACGAAGAAATTATTAATCTGCTGGAAGAAATATCTGGGATTATTCAGATTTGTA 4888
 Qy 567 -----GlnSerThrValSerIlybGlnAspAlaMetProGln-----AlaLeuIlybSer 582
 Db 4889 CAAGGGAAACAAATATTAAGACCTTGATGATATTCACAGAAATACACATTTAAGAAA 4948
 Qy 583 LeuIlePheProAsnPheG1ProLeuHisIlybAspHisIthrAsnPheLeuIlyb----- 600
 Db 4949 GTAGTTTTCGAAGTAC-----CTGACAGCAGCACTTTTACCAAAACGTGAT 4993
 Qy 601 -----GluIleGlnIleValGlnAlaLeuThrGlnIleIlybSerAsnAlaGlnIle 617
 Db 4994 CTTGAGGGAGATGATGAAGACATCTCGAAATCTTGCTCTTATTCAGAGCTGAAAGTG 5053
 Qy 618 ArgAspGlyroGlnArgIleGlyAspValMetLeuIlybAsnIleGlnIlybMetIlybHis--- 636
 Db 5054 GATGAT-----GTTATCCAAAGTTCGTTTAAATTTACATCCAGACCGTATTTCT 5104
 Qy 637 LeuAlaAlaHisIleuThrIlybHisSerGlnAlaLeuGlnIleuGlnIleuGlnIleIlyb 656
 Db 5105 CTGATGTCCTCCATGAGCTTCTCTGTAAGAAATATCAAGAGGAAAGACCTTATGAAGATGAC 5164
 Qy 657 SerSerArgArgLeuGlnIlybAsnPheCysValAspPheGlnIleuGlnIlybValCysGlyIleu 676
 Db 5165 TCTCTCT-----GATAGGTTTAAAGAAATATGCTGCTCAG----- 5200
 Qy 677 ProLeuAsnThrPheLeuIleuArgProLeuHisIlybLeuMetHisIlybIlybGlnValLeu 696
 Db 5201 -----CDAATAGCAATTATCTTAAGATGAGCTTAC 5236
 Qy 697 GluArgIleuCysIlybHisIlybProSerHisIleAspPheArgAspCysValGlyAla--- 715
 Db 5237 GAGCGATCAGGGCAGAAAGTCTGGCTCT-----GCCTGCGGGCTTGG 5281
 Qy 716 AlaLeuAlaGlnIleThrGlnMetValAlaGlnLeuHisIleGlyThrMetIleIlybMetGln 735
 Db 5282 GCTTACTACTCAGCTGTAACCCAGAAAGATATTACTCAGAGCTTGATGAGTCCCTC 5341
 Qy 736 AsnPheGlnIlybLeuHisIleGlnIlybIlybAspLeuIleGlyIleAspAsnLeuValI 755
 Db 5342 CTATCTGCCAAAGGAATACTACATAGAGC-----AGCAACATGCTGTG 5389
 Qy 756 ProGlyArgGlnPheIleArgLeuGlnIlybSerIleuSerIlybLeuSerGlnIlybGln 775
 Db 5390 GACATGAGAAATTTGCTTACTTA----- 5413
 Qy 776 GlnArgMetPhePheLeuPheAsnAspValLeuLeuTyr----- 788
 Db 5414 -----TTTATGACCTTTGATGAAGAAAGAGAGCTGATATGAAAGAACTCCGA 5464
 Qy 789 ThrSerArgIlybLeuThrAlaSerAsnGlnPheIlybValHisIleGlnIleuProLeuTyr 808
 Db 5465 AGTTCCTGAGACTACATCTCAGAAATGAAGAAATTCATAGATTTCAAGCCATTAAT 5524
 Qy 809 GlyMetThr-----IleGlnIleuSerGlnIlybAspIlybThrGlyValIleProHis 823
 Db 5525 GGGATGATCTAAGCTTGGCTGCTGGAAGAAATGATAGAGAAAGTGGGAGGTTTAC 5584
 Qy 824 CysLeuThrIleuArgGlnIleArgGlnSerIleIleValAlaAlaSerSerArgSerGln 843
 Db 5585 CTGCTGAACATTAAGCGCTCCCAAGACATTCACGTGGTTCATCTTCAAGCCAGACCTTG 5644
 Qy 844 MetGlnIlybThrValGlnIleAspIleGlnMetAlaIleAspLeuAlaGlnIlybSerSerSer 863
 Db 5645 CTGAAATAGGCATTAACAGCAGCAAGTTAAGGGGTCTGGCCCTTCTGCTGGATTCAATT 5704

Oy 864 ProAlaProGluPheLeuAlaSerSerPro 873
|||
Db 5705 AAAACTCTTGAAATGAGGACATCAAAACT 5734
|||

Search completed: May 20, 2006, 06:59:15
Job time : 795 secs

This Page Blank (uspto)